Identification of Differentially Expressed Genes in Normal and Malignant Prostate by Electronic Profiling of Expressed Sequence Tags

Yan W. Asmann,1 Farhad Kosari,1 Kai Wang, John C. Cheville, and George Vasmatzis2

Division of Experimental Pathology, Department of Laboratory Medicine and Pathology, and Mayo Cancer Center, Mayo Clinic and Foundation, Rochester, Minnesota, 55905

ABSTRACT

Differentially expressed genes between corresponding normal and cancer tissue can advance our understanding of the molecular basis of malignancy and potentially serve as biomarkers or prognostic markers of malignancy. To identify differentially expressed genes in prostate cancer, we used a procedure combining electronic expression profiling of the prostate expressed sequence tag (EST) database and molecular biology techniques. A novel electronic expression-profiling algorithm was developed to search publicly available EST sequences for genes that show significant differential expression in prostate cancer compared with normal prostate tissue. Approximately 600 genes expressed in prostate were identified through adequate EST counts of ESTs for electronic profiling. Of these 600 genes, 9 showed statistically significant differences in their EST counts between cancer and normal prostate and were further analyzed. The predictions associated with electronic profiling were experimentally verified for two genes, cysteine-rich secretory protein 3 (CRISP-3) and deadenylating nuclease (DAN), using real-time reverse transcription-PCR with total RNA extracted from cells isolated by laser capture microdissection. In five of five Gleason score 6 cancer cases, CRISP-3 expression was increased >50 fold, whereas the expression of DAN was reduced by >80%.

INTRODUCTION

Cancer-associated changes in gene expression are now routinely investigated by high throughput expression profiling techniques such as SAGE3 (1), microarray/gene chips (2), and sequencing of clones from cDNA libraries (3, 4). An important consideration in expression profiling is the quantity and purity of RNA obtained from the different cell types. Increasing the quantity and purity of RNA from either malignant or benign epithelia invariably yields more reliable and comprehensive experimental results. Although LCM has helped with regard to this issue (5), cell selection and extraction by LCM is currently a manual process, and at most only a few thousand cells can be extracted in an amount of time necessary to limit RNA degradation.

Electronic analysis of sequenced cDNAs uses the information in the EST database (dbEST) where >3.5 million sequences have been deposited from the analysis of over 6,000 human cancer and normal cDNA libraries constructed from >50 different tissues (6).4 To use this large amount of information, computer algorithms have been developed for discovery of novel genes (7) and genes with limited tissue distribution and/or cancer-specific expression (8). One limitation associated with the use of this dbEST is that only highly expressed genes have been sampled adequately to provide sufficient corresponding EST counts for reliable molecular profiling. However, identification of these highly expressed genes could provide significant information to enhance our understanding of carcinogenesis and serve as biomarkers or prognostic markers of malignancy. Prostate is an excellent tissue to study by EST analysis because of the sizable pool of EST data in the LCM-derived libraries.

Despite the existing search algorithms, there is still a need for sophisticated computer analysis tools to perform the clustering and analysis of the ESTs. In this study, we developed a novel electronic profiling algorithm, the Binary Indexing Search Algorithm, to identify differentially expressed genes in cancer and normal prostate EST databases. Functions of this novel algorithm include clustering ESTs to distinct genes, estimating tissue distribution of EST clusters, and sorting EST clusters by relative expression levels. This procedure, using the Binary Indexing Search Algorithm, requires ~6 h of CPU time, and in this study, we successfully used this algorithm for the identification of genes in cancer and normal microdissected prostate EST libraries. The tumor-associated differential expression for two genes, CRISP-3 and DAN, was subsequently verified experimentally.

MATERIALS AND METHODS

Binary Indexing Search Algorithm

Classification of dbEST. The dbEST was categorized using the GLS tool of the cancer genome anatomy project.5 We considered precancer libraries as cancer libraries. All of the cDNA libraries were categorized according to tissue type (tissue origin), tissue histology (cancerous, normal, or fetal), and the library preparation method (microdissected, bulk, cell line, or flow cytometric sorted).

Data Preparation. The data preparation module reads EST report files downloaded from NCBI and produces files in FASTA format for ESTs.

Tissue Distribution of ESTs/Genes. In the tissue distribution module, binary indexes were used to assign tissue distribution to each prostate EST. A binary index is defined as a binary string of n bits, where n is an integer. A sequence of n nucleotides can be represented by n-bit binary index when purines (As and Gs) are assigned to 1 and pyrimidines (Ts and Cs) are assigned to 0. After a series of testing, we found that a 32-bit binary index is long enough to rule out a random occurrence, such that under most circumstances a unique 32-bit index represents a unique EST. In this study, we used 32-bit continuous indexes to represent ESTs, and every EST was converted into groups of continuous 32-bit binary indexes.

In the first step, the program declares, in random access memory, a DA table for all possible indexes, in which each entry has one byte (8 bits) of memory reservation. The first distinct 29 bits of a 32-bit index are used as the DA for that index in the table, and the decimal value (d_j) of the last three bits will give the bit position to be set to 1 in that entry. For example, a 32-bit index, 11110011110111101111100100011111, has a direct entry address 11110011110111101111100100011111 in the table, and the 7th bit of that entry is set to 1 [d_j = 2 (2) + 2 (1) + 2 (0) = 7]. We ran this module on classified prostate dbEST and created a complete nonredundant DA table containing all prostate indexes.

The algorithm then sets up a tissue distribution table (OS table) using the DA table. Each entry of the OS table contains two fields: a 32-bit index and a 64-bit specificity string (there are 52 different tissue types in human dbEST). This module runs on the entire human dbEST to get every possible index and corresponding tissue.

Received 10/2/01; accepted 4/5/02.

1 These authors contributed equally to this work.

2 To whom requests for reprints should be addressed, Mayo Clinic and Foundation, 200 First Street, SW, Rochester, MN 55905. Phone: (507) 266-4617; Fax: (507) 266-5193; E-mail: vasmatzis.george@mayo.edu.

3 The abbreviations used are: SAGE, serial analysis of gene expression; EST, expressed sequence tag; LCM, laser capture microdissection; GLS, Gene Library Summarizer; DA, direct addressing; RT-PCR, reverse transcription-PCR; GAPDH, glyceraldehyde-3-phosphate dehydrogenase; NM, normal microdissected; CM, cancer microdissected; CB, cancer bulk; NB, normal bulk; CRISP-3, cysteine-rich secretory protein 3; TARP, T-cell receptor gamma transcript; PSA, prostate-specific antigen.


5 Internet address: http://cgap.nci.nih.gov/Tissues/LibrarySummarizer.

check if each index is a prostate index using the DA table. If an index is a prostate index, the module builds an entry of this index in the OS table as follows. The direct address of this entry will be the first 24 bits of the index (we use 24-bit addresses to reduce the table size). In that address, we store the complete 32-bit index and the 64-bit specificity string of which the bit corresponding to the tissue where this index comes from will be set to 1. Because we used external chaining to store the indexes that have identical first 24-bits but different last 8 bits, the OS table is also a nonredundant table. After running the human dbEST on this module, each entry in the DA table has a cumulative tissue distribution string in OS table showing all of the tissues corresponding to the index. Therefore, the tissue distribution of each prostate EST can be derived from the OS table by accumulating the tissue distribution of all of the indexes related to that EST. For the same procedure, the tissue distribution of an EST cluster (see below) or a gene can be calculated by accumulating the tissue distributions of all of the ESTs in that cluster.

**EST Clustering.** To cluster the ESTs belonging to the same gene together, we created a clustering module in our binary indexing search algorithm. All of the prostate ESTs in FASTA formatted files generated by data preparation module are read. Each EST was then converted into groups of continuous 32-bit binary indexes. For every individual EST, if none of its indexes pointed to any previously built cluster, the EST and all its indexes are assigned to a new cluster. Otherwise, if any of the indexes of an EST pointed to an already built cluster, this EST would be assigned to that cluster, and all its indexes would be set to point to the same cluster.

**Sorting and Selecting EST Clusters.** Within each prostate EST cluster, we grouped ESTs according to the tissue histology (cancerous or normal) and the library preparation method (microdissected, bulk, or cell line). This resulted in five groups of ESTs including NB, NM, CB, CM, and CL. Then, we sorted clusters according to the total EST counts, yielding information regarding gene expression level in prostate, differential expression level of the same gene between normal and cancer. Furthermore, we selected the EST clusters with statistically significant EST counts and/or the statistically significant distribution of an EST cluster (see below) or a gene can be calculated by plotting the histogram of all of the indexes related to that prostate gene. There are few ESTs, this plot provides an estimate of the differential expression for clusters with very small EST counts.

**Experimental Procedures**

**Tissue Section and H&E Staining.** Five cases of Gleason score 6 OCT-embedded prostate frozen tissues as well as their corresponding normal tissues were used. Ten-μm cryostat sections were stained using H&E and dehydrated completely in xylene, following protocols from Arcturus Engineering (Mountain View, CA).

**LCM.** The LCM technique allows for isolating normal and cancerous prostate epithelial cells precisely and efficiently from among a mixture of normal, cancer, epithelial, and nonepithelial cells. Navigated-LCM was performed with a ProCell II apparatus (Arcturus) on H&E sections. LCM parameters included a laser power of 65 milliwatts, laser pulse duration of 1.2 ms, and laser spot size of 7.5–15 μm in diameter. The infrared laser was pulsed over cells of interest. We captured approximately the same number of cancerous and normal epithelial cells (~2,000) in each case.

**RT-PCR.** After LCM, total RNA extraction from the captured cells was performed using the RNaseasy kit (Qiagen). The LCM-captured cells were immediately placed into sterile 0.5-ml microcentrifuge tubes containing 200 μl of RLT reagent with 1% β-mercaptoethanol and inverted at room temperature for 1 hour before extraction of total RNA. The RNA extraction was performed according to the instructions from RNeasy kit. Total RNA was eluted into 30 μl of nuclease-free H2O. Oligo-dT primers and SuperScriptII reverse transcriptase (Life Technologies, Inc.) were used in reverse transcription. The reverse transcriptase was performed following the manufacturer’s protocol. After the reverse transcription reaction, 40 rounds of PCR amplification were performed using gene-specific primers. The gene-specific forward and reverse primers were designed using the Primer 3 program.

PCR amplifications were performed in a total volume of 20 μl, with 1 μl of the reverse transcription products, 1 unit of Taq polymerase, 0.5 μM of each primer, 0.25 mM of each deoxynucleotide triphosphate in 1× PCR buffer.

**Real-Time RT-PCR.** To obtain a more accurate estimate of the changes in mRNA expression levels in cancerous prostate epithelial cells versus normal/benign prostate epithelial cells, real-time RT-PCR was used. We used the RT product from RT-PCR as templates in this experiment, together with the TaqMan universal PCR MasterMix (Applied Biosystems), gene-specific primers, and dual-labeled TaqMan probes designed using the Primer Express 1.5 software (PE Biosystems). The highly gene-specific regions that do not display any sequence homology with their close family members were chosen for amplification.

**RESULTS**

**Development of a Binary Indexing Search Algorithm to Analyze Prostate ESTs.** All of the prostate cDNA libraries were categorized using GLS tools according to tissue histology (cancerous or normal) and the library preparation method (microdissected, bulk, or cell line). The Binary Indexing Search Algorithm was used to analyze the classified EST data. The Binary Indexing Search Algorithm contains four modules: a data preparation module for converting the publicly available EST information to FASTA formatted files; a tissue distribution module for assigning tissue distributions to each index/EST/gene; an EST clustering module for clustering ESTs by different genes; and a selection module for sorting EST clusters according to the number of ESTs in each cluster, and selecting the genes of interest. The data flow between these four modules is shown in Fig. 1.

**Identification of Differentially Expressed Genes between Normal and Cancer Prostate Using Binary Indexing Search Algorithm.** To date, human prostate EST libraries have >130,000 prostate ESTs, including ESTs from NB, NM, CB, CM, and CL libraries. The prostate EST database is the fifth largest EST database after colon, kidney, lung, and breast libraries, having EST entries from close to 300 prostate libraries. We postulate that, given the size of prostate cDNA libraries, the gene expression level in prostate correlates with the number of ESTs in the corresponding gene cluster.

Prostate adenocarcinoma is of epithelial origin. Therefore, we confined our analysis to CM and NM pools, which contain 22,776 and 9,967 ESTs, respectively, to avoid erroneous calculation of expression levels attributable to contamination from nonepithelial cells. After analyzing these two EST pools using the Binary Indexing Search Algorithm, the expression profile of close to 600 EST clusters in prostate CM and NM were identified. Fig. 2 displays a plot that represents the distribution of gene regulation in prostate for these 600 genes. EST clusters (distinct genes) were grouped according to a pseudo-ratio (r) of normalized numbers of ESTs in NM and CM pools, as follows:

\[
r_c = \text{INT} \left( \frac{(cm/MAX \ (nm, 1)) - 1}{(cm, 1)} \right)
\]

for up-regulated genes, where \(nnm = nm\ (CM/NM)\).

\[
r_d = \text{INT} \left( \frac{(mmn/MAX \ (cm, 1)) - 1}{(cm, 1)} \right)
\]

for down-regulated genes (\(nnm\), the number of gene-related ESTs from NM; \(cm\), the number of gene-related ESTs from CM; \(cm\) and \(NM\) represent the total number of ESTs in CM and NM libraries; MAX, the function to return the maximum value of input parameters; INT, the function to return the integer value of input parameter).

To avoid infinite values for \(r\), the minimum value of the denominators was 1. The \(r\) value for down-regulated genes was negated to separate down-regulated genes from the up-regulated genes in the plot. The number 1 was subtracted from the quotient so that \(r\) for nonregulated genes was assigned to 0. The distribution curve of prostate gene regulation can be illustrated by plotting the histogram of \(r\) values versus the number of clusters (Fig. 2). Although \(r\) does not precisely describe the differential expression for clusters with very few ESTs, this plot provides an estimate of the differential expression of genes between normal and cancerous prostate epithelium. As shown in Fig. 2, the largest group of EST clusters consists of genes that equally express between normal and prostate cancer. There are
The T-cell receptor with prostate cancer. TARP is a widely used housekeeping gene (GAPDH), a yeast transcription activator. The function of GAPDH is not clear. The DANCE gene encodes for a secretory protein that has been shown to promote adhesion of endothelial cells (12). The TARP gene is a member of the family of actin filament-binding proteins and has been reported to be down-regulated in prostate cancer (13). NPD017 is a novel gene without a putative ID in dbEST. The finding of semenogelin II as a down-regulated gene in prostate cancer is unexpected. Semenogelins are the predominant proteins in human semen and are the major proteins involved in the gelatinous entrapment of ejaculated spermatozoa (14). They are synthesized by the secretory epithelium of the seminal vesicles and should not be present in prostate microdissected EST libraries. It is possible that semenogelin II ESTs in prostate NM and CM libraries are the result of contamination of seminal vesicle cells in prostate tissue cDNA samples. In fact, when we carefully selected only cancer or normal prostate epithelial cells by LCM and performed RT-PCR, the semenogelin II band was absent, even after 40 cycles of the PCR reaction (Fig. 4d). Table 1 also includes several genes from the expression distribution curve with no differential expression to use as controls for normalization. PSA, HK2, and PAP are prostate epithelium-specific genes, and GAPDH is a widely used housekeeping gene expressed in all cells. Although serum concentrations of PSA and HK2 are elevated in most prostate cancer patients (15), our analysis in a few

close to 160 genes in this group including prostate-specific genes such as the PSA (KLK3 and PSA), prostate kallikrein 2 (KLK2 and HK2), prostatespecific acid phosphatase (PAP), and the housekeeping gene GAPDH. The most up-regulated gene in prostate cancer identified by our algorithm was CRISP-3, with an r of 32. The group of down-regulated genes, including a tumor suppressor gene RT14, had r values as low as ≲23. According to Fisher’s exact test, there had to be at least 8 ESTs in the CM pool for up-regulated genes and at least 3 ESTs in NM pool for down-regulated genes to be a statistically significant difference between CM and NM. Thirty-three of 600 genes that passed the Fisher’s exact test and the 1-tail test are listed in Table 1. Additionally, we also included differentially expressed genes that had an r ≥3 even though they failed the Fisher’s exact test and the 1-tail test. The complete list of these 600 genes is available on Internet address: http://www.mayo.edu/research/expath/prostate.html.

Table 1 also includes several genes from the expression distribution curve with no differential expression to use as controls for normalization. PSA, HK2, and PAP are prostate epithelium-specific genes, and GAPDH is a widely used housekeeping gene expressed in all cells. Although serum concentrations of PSA and HK2 are elevated in most prostate cancer patients (15), our analysis in a few

![Data flow within the Binary Indexing Search Algorithm. The arrow symbolizes the data direction. Rectangles with dotted borders indicate different modules. Raw EST sequences were imported from prostate EST libraries and classified using the publicly available GLS tool. Different EST pools were processed through the Binary Indexing Search Algorithm. Each periodic result (as indicated in rectangles) was saved and directly accessible.](image-url)
samples shows that the per cell mRNA expression levels of PSA and HK2 do not change between normal and malignant prostatic epithelia.

### Relative Expression Level of 9 Genes in Normal and Cancerous Prostate

**Table 1** Differentially displayed genes in normal and cancerous prostate

| Gene name OR GenBank ID | CRISP-3 | PSA, Hk2, PAP, and GAPDH, and PAP as standards (Fig. 3). We found that PSA is the highest expressed gene in both prostate cancer and normal epithelial cells, which represents about 1.3% of total ESTs in both microdissected libraries (CM and NM). In addition, from our electronic profiling, CRISP-3 is relatively prostate specific (CRISP-3 is found in four organs; Table 1) and highly expressed in prostate cancer (0.15% of total expressed cancer transcripts; Fig. 3). The putative high expression level of CRISP-3 in cancer, which is higher than or as least as high as that of GAPDH, and its secretory feature make it a potential candidate for a highly sensitive and specific diagnostic marker for prostate cancer.

**RT-PCR Confirms the Expression Regulation of CRISP-3 and DAN in Primary Prostate Cancer.** We examined the putative up-regulation of CRISP-3 and the putative down-regulation of DAN in prostate cancer using RT-PCR. PSA, HK2, PAP, and GAPDH were used as standards. Total RNAs from ~2,000 LCM-captured prostate epithelial cells (Gleason score 6 and the corresponding normal/benign) were used in RT-PCR (Fig. 4, a–c). As shown in Fig. 4d, CRISP-3 and DAN exhibit different expression patterns between normal and cancer, whereas standards PSA, HK2, PAP, and GAPDH show...
approximately equal expression levels in cancer and normal. We investigated 5 Gleason Score 6 prostate cancer cases and obtained similar results in all 5 cases.

**Quantification of Differential Expression of CRISP-3 and DAN Using Real-Time RT-PCR.** To obtain a quantitative estimate of expression levels of the above 6 genes, real-time PCRs was performed next using TaqMan PE 7700 system (Applied Biosystems). The TaqMan assay uses the 5′→3′ exonuclease activity of Taq DNA polymerase and a fluorogenic probe for automated quantification of DNA in a real-time manner. The CT value refers to the threshold cycle at which a statistically significant increase in fluorescence is first detected by the sequence detection system. The increase in fluorescence is directly proportional to the exponential increase in PCR products, and the measurement of signal is carried out in a real-time manner.

Because of the possible different amplification efficiencies of different genes, a validation experiment to calculate individual amplification efficiency for each gene was performed. The total RNA from the prostate cancer cell line LNCaP, which contains mRNA from all 6 genes, was used as the starting material. The CT values were obtained from starting total RNA of 10, 5, 2.5, 1.25, 0.625, and 0.3125 ng for each gene. The amplification efficiency of each gene was calculated following the manufacturer’s procedure (Table 2). The differential expression of a gene in cancer is computed as $e^{\Delta C_{T}}$, where $e$ is the amplification efficiency and $nC_{T}$ and $cC_{T}$ denote the CT for normal and cancer, respectively. As shown in the table, GAPDH, PSA, H2, and PAP are approximately equally expressed in cancer and normal prostate epithelial cells. CRISP-3 is significantly up-regulated (50–300-fold) in prostate cancer, whereas the tumor suppressor gene DAN is down-regulated 84–94% in prostate cancer when compared with normal prostate epithelium.

**Comparison of the Binary Indexing Search Algorithm Profiling and Other Profiling Methods.** Some differentially expressed genes are identified by other profiling methods including microarray, GeneChip, and SAGE that were not present in our profiling results (Table 3). Among these, hepsin has been consistently reported as a significantly up-regulated gene in prostate cancer using microarray and GeneChip analyses (16). However, our study showed that there are only 3 hepsin ESTs in the prostate CM pool, and hepsin did not pass the Fisher’s exact test. Cytokeratins 8, 18, and 19 are reported down-regulated in high grade cancer (Gleason score 8–9; Ref. 17), whereas the r value according to our calculation is 0. This may be because the CM pool used in this study does not distinguish between cancer grades. Other genes in Table 3 such as glutathione S-transferases (18), α-methyl-CoA racemase (19), KIAA 1538 protein (20), human tyrosine kinase receptor axl (21), and PSGR (22) were not present in our data sets because they did not pass the Fisher’s exact test or did not have ESTs in the CM and NM pool.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Amplification efficiency ($e^a$)</th>
<th>ΔCT $^b$</th>
<th>Expression regulation (cancer/normal)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GAPDH</td>
<td>1.99 ± 0.02</td>
<td>0.11 ± 0.3</td>
<td>0.88–1.32</td>
</tr>
<tr>
<td>PSA</td>
<td>1.94 ± 0.10</td>
<td>0.00 ± 0.0</td>
<td>1.00</td>
</tr>
<tr>
<td>H2</td>
<td>1.87 ± 0.02</td>
<td>0.11 ± 0.5</td>
<td>0.68–1.78</td>
</tr>
<tr>
<td>PAP</td>
<td>1.82 ± 0.02</td>
<td>-1.88 ± 1.5</td>
<td>0.05–2.07</td>
</tr>
<tr>
<td>CRISP-3</td>
<td>1.94 ± 0.1</td>
<td>7.29 ± 0.8</td>
<td>49.6–316.95</td>
</tr>
<tr>
<td>DAN</td>
<td>1.92 ± 0.03</td>
<td>-3.63 ± 0.8</td>
<td>0.06–0.16</td>
</tr>
</tbody>
</table>

$a$ The efficiency for linear amplification is 2.

$b C_{T_{normal}} - C_{T_{cancer}}$
We compared the lists of differentially expressed genes identified by other profiling methods to our analysis of the dbEST (Table 1). None of the four characterized most up-regulated genes identified by our profiling method, CRISP-3, 6.2kd protein (LOC54543), KIAA1666 protein, and fatty acid binding protein 5, or the 4 characterized most down-regulated genes NPD017, creatine kinase-like gene, UBIQ, and hypothetical protein FLJ211174 were identified using other profiling methods.

Microarray and GeneChip contain a limited probe set, usually 5,000–6,000 known genes. We found that, except for the hypothetical protein FLJ211174, none of the other 8 up-regulated genes identified by our analysis were included in most microarray/GeneChip probe sets. In addition, microarray, GeneChip, and other profiling technologies usually require relatively large quantities of total RNA. As a consequence, bulk tissues (normal and cancerous) instead of LCM-captured cells are routinely used in the analysis. These profiling methods may identify genes that are not specific to normal or malignant prostatic epithelia. This may also add to the discrepancy in genes identified in our analysis when compared with other gene profiling methods. As result, the Binary Indexing Search Algorithm offers additional insight into gene expression that other profiling studies may not be able to provide.

**DISCUSSION**

In the past decade, an effort to sequence expressed genes in different tissues has made analysis of the dbEST, an alternative resource for gene expression profile analysis. Our study combined analysis of the dbEST using a novel Binary Indexing Search Algorithm with traditional molecular biology techniques to identify genes that are up- and down-regulated in prostate cancer. Using the Binary Indexing Search Algorithm, we performed analysis on the human prostate EST database and identified a list of tumor-regulated genes in prostate cancer. Compared with traditional search algorithms, the Binary Indexing Search Algorithm uses direct addressing to distribute indexes and EST data in memory. Therefore, it avoids \( n^2 \) comparisons between sequences and is able to quickly identify tissue specificity and cluster ESTs. Although it demands moderately large memory capacity, it uses only a few hours of CPU time to perform the entire procedure. In addition, our approach takes advantage of the publicly available prostate cDNA libraries. Therefore, unlike other approaches such as SAGE or DNA microarrays, the laborious experimental step before electronic profiling is not necessary.

Some consideration should be taken while viewing the output of the Binary Indexing Search Algorithm. For instance, because of the small size of some EST libraries from different organs or tissue types, there may be more organs or tissue types where the gene is expressed than identified by the algorithm. For example, CRISP-3 was found in four organs according to the computational calculation. However, CRISP-3 was found in seven different organs by multiple tissue dot blot including salivary gland, pancreas, prostate, ovary, thymus, testis, and colon.

Prostate cancer is the most commonly diagnosed noncutaneous malignancy and the second leading cause of cancer-related deaths in the Western male population (23). Currently, measurement of the serum PSA is the most sensitive biomarker for the detection of prostatic adenocarcinoma. However, an elevation in the serum PSA lacks specificity, and serum PSA may be elevated in common benign conditions of the prostate such as prostatitis and benign prostatic hypertrophy. Numerous studies have shown that only 25% of patients with an elevated serum PSA level between 4 and 10 ng/ml have an adenocarcinoma detected on prostate needle biopsy (24, 25). The lack of specificity for PSA results in unnecessary prostate needle biopsy procedures and patient anxiety. More specific biomarkers for prostate cancer are needed to improve our ability to detect prostate cancer.

Many studies, as well as our electronic profiling results, have shown that prostate cancer expression levels of PSA on a per cell basis do not change and may even decrease in high-grade prostate cancers.

Our data demonstrate that the per-cell CRISP-3 mRNA level is significantly up-regulated in prostate cancer compared with normal tissue. In addition, there is evidence to indicate that CRISP-3 is a secretory protein. Therefore, CRISP-3 is a potential diagnostic marker for prostate cancer, and our subsequent studies will focus on the utility of CRISP-3 as a diagnostic marker for prostate cancer. Similar to PSA, CRISP-3 is also androgen responsive (26), although the functions of CRISP-3 and its close family members remain largely unknown.

**REFERENCES**

10. Unpublished data.
Identification of Differentially Expressed Genes in Normal and Malignant Prostate by Electronic Profiling of Expressed Sequence Tags


Updated version
Access the most recent version of this article at:
http://cancerres.aacrjournals.org/content/62/11/3308

Cited articles
This article cites 23 articles, 11 of which you can access for free at:
http://cancerres.aacrjournals.org/content/62/11/3308.full.html#ref-list-1

Citing articles
This article has been cited by 8 HighWire-hosted articles. Access the articles at:
/content/62/11/3308.full.html#related-urls

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, contact the AACR Publications Department at permissions@aacr.org.