Neoplasia • Vol. 8, No. 3, March 2006, pp. 181–189 181

www.neoplasia.com

A Novel Method for Gene Expression Mapping of Metastatic Competence in Human Bladder Cancer[1]

Z. Wu[*,2], *M. S. Siadaty*[†,2], *G. Riddick*[†,2], *H. F. Frierson Jr.*[‡,2], *J. K. Lee*[†,2], *W. Golden*[‡,2], *S. Knuutila*[§,2], *G. M. Hampton*[¶,2], *W. El-Rifai*[#,2] *and D. Theodorescu*[*,2]

*Department of Molecular Physiology, Box 422, University of Virginia Health Sciences Center, Charlottesville, VA 22908, USA; [†]Division of Biostatistics, Department of Public Health Sciences, Box 674, University of Virginia Health Sciences Center, Charlottesville, VA 22908, USA; [‡]Department of Pathology, Box 800214, University of Virginia Health Sciences Center, Charlottesville, VA 22908, USA; [§]Department of Pathology and Molecular Cytogenetics, Haartman Institute, University of Helsinki, Helsinki, Finland; [¶]Genomics Institute of the Novartis Research Foundation, 10675 John Jay Hopkins Drive, San Diego, CA 92121, USA; [#]Department of Surgery and Ingram-Vanderbilt Cancer Center, Vanderbilt University Medical Center, Nashville, TN, USA

Abstract

Expression profiling by DNA microarray analysis has provided insights into molecular alterations that underpin cancer progression and metastasis. Although differential expression of microarray-defined probes can be related to numerical or structural chromosomal alterations, it is unclear if such changes are also clustered in distinct chromosomes or genomic regions and whether chromosomal alterations always reflect changes in gene expression. Here we apply the dChip algorithm and a novel technique to test the hypothesis that expression changes occurring as a function of tumor progression and metastasis are nonrandomly distributed. Expression profiling of a human xenograft model of lung metastasis phenotype indicates that chromosomes 2, 11, and 20 contain higher percentages of differentially expressed genes ($P < .05$). Furthermore, we show that a number of differentially expressed probes mapped to chromosome 17q, defining the existence of an expression ''hot spot'' corresponding to an area of gain determined by comparative genomic hybridization (CGH). Interestingly, other areas of gains detected by CGH were not associated with expression hot spots. In summary, we show that gene expression changes during bladder cancer lung metastasis occur nonrandomly in specific chromosomes and intrachromosomal locations.

Neoplasia (2006) **8,** 181–189

Keywords: Bladder neoplasms, metastasis, gene expression, comparative genomic hybridization, karyotype.

Introduction

The use of DNA microarrays has revolutionized the measurement of gene expression in human cancer by providing insights into genes involved in tumorigenesis and metastasis [1], into the discovery of potential tumor biomarkers [2,3], and into the molecular classification of common neoplasms [4]. Some differentially expressed genes have been associated with specific chromosomal gains and losses, as determined by karyotyping and comparative genomic hybridization (CGH) [5,6], whereas others appear not to be due to structural chromosomal alterations. Thus far, no evaluation has yet been performed to determine whether such gene expression changes occur either in specific chromosomes or in distinct chromosomal regions (''hot spots'') during bladder tumor metastasis.

In the current study, we use both an established and a novel bioinformatics approach to the discovery and mapping of gene expression hot spots. We apply these approaches to gene expression profiling data obtained from a newly described model of lung metastasis in progressive bladder cancer [7], and we use this model to determine if genes whose expression is altered as a function of increasing metastatic potential are randomly distributed or clustered in specific chromosomes or chromosomal regions. In addition, we evaluate the bladder cell lines in this model for genomic alterations to see whether such alterations are associated with gene expression hot spots.

Materials and Methods

Xenograft Model of Lung Metastasis in Progressive Bladder Cancer

The xenograft model of lung metastasis in progressive bladder cancer was recently described [7]. In this model, a

Address all correspondence to: Dan Theodorescu, Department of Molecular Physiology, Box 422, University of Virginia Health Sciences Center, Charlottesville, VA 22908.
E-mail: dt9d@virginia.edu

[1]This work was supported by National Institutes of Health grant CA075115.
[2]The authors declare that they have no conflicting financial interests.
Received 3 November 2005; Revised 17 December 2005; Accepted 19 December 2005.

Copyright © 2006 Neoplasia Press, Inc. All rights reserved 1522-8002/06/$25.00
DOI 10.1593/neo.05727

sequential series of cell lines was generated from a set of more progressive metastatic tumors selected *in vivo*. To accomplish this, T24T cells [8], which are tumorigenic and poorly metastatic variants of T24, were serially passaged in nude mice. The derivative cell lines with increasing pulmonary metastatic ability were named FL1, FL2, and FL3 (Figure 1). Gene expression analysis of T24T, FL1, FL2, and FL3 was performed as previously described using the HG-U133A GeneChip array (Affymetrix, Santa Clara, CA). Upregulated and downregulated genes were identified as a function of metastatic ability [7] and were used in the gene mapping analysis described below.

Karyotypic Analysis

Cultures of the T24T cell line and the sublines FL1, FL2, and FL3 were initiated by placing aliquots of trypsinized cells on small glass coverslips. The coverslips were incubated for 4 to 7 days at 37°C in minimum essential medium (Gibco, Grand Island, NY) supplemented with 20% fetal bovine serum and penicillin/streptomycin (Gibco). When adequate mitotic cells had been observed, the coverslips were treated with Colcemid (Gibco) overnight and harvested *in situ*, according to standard methods [9]. The coverslips were stained using a standard trypsin–Giemsa banding method. Five karyotypes were produced from each cell line. The final karyotype designation was based on a composite, as indicated by "cp" at the end of the description and following standard ISCN 1995 nomenclature.

CGH

DNA was extracted from the T24T and FL1–FL3 cell lines, and CGH was performed using a mixture of fluorochromes conjugated to dCTP and dUTP nucleotides for nick translation [10]. Hybridization, washing, and ISIS digital image analysis (Metasystems GmbH, Altlussheim, Germany) were performed as described [11]. All CGH results were confirmed using a 99% confidence interval. Briefly, intraexperiment standard deviations for all positions in CGH ratio profiles were calculated from the variation of ratio values of all homologous chromosomes within the experiment. Confidence intervals for the ratio profiles were then computed by combining them with an empirical interexperiment standard deviation and by estimating error probabilities based on *t* distribution. For the analysis of the frequencies of DNA copy number changes, we accepted only changes seen using fixed cutoff values and confirmed with 99% confidence.

Controls

In each CGH experiment, a negative control (peripheral blood DNA from a healthy donor) and a positive control were included. The positive control was a gastric tumor with known DNA copy number changes. Based on our earlier reports and on control results, we used 1.17 and 0.85 as cutoff levels for gains and losses, respectively. High-level amplification (HLA) was considered at $\geq$1.50.

Chromosome Mapping of Genes Differentially Expressed in Association with Tumor Progression

For the discovery of hot spots, we applied two different techniques. The first is the "Genome View" algorithm from dChip, a popular program used in the analysis of gene expression data (http://www.dchip.org/). The second is a novel method developed for this manuscript based on a

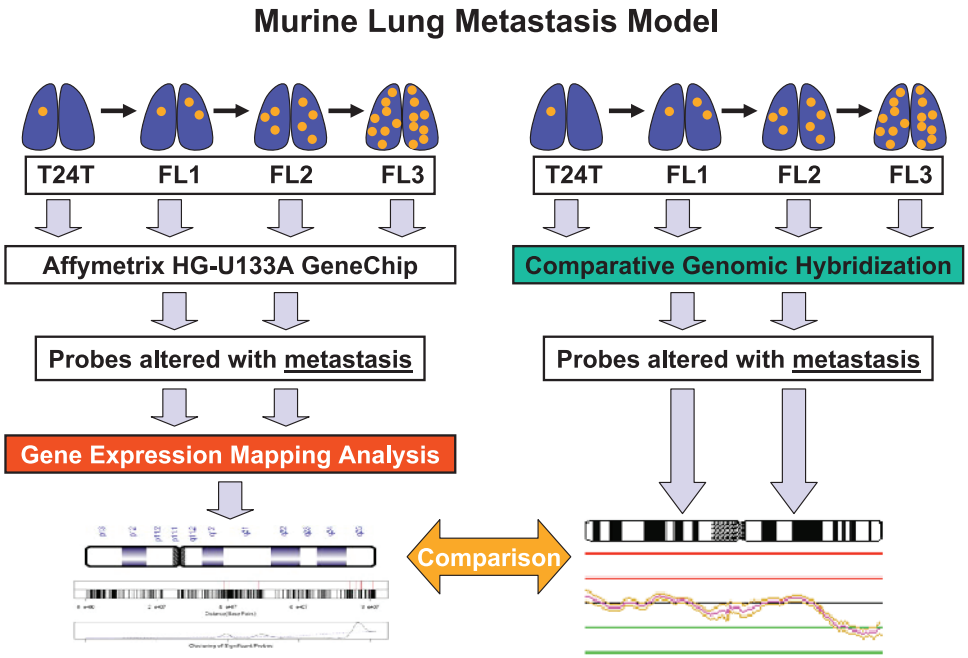

Figure 1. *Diagram of model and analysis. A diagram of cell lines with increasing lung metastatic ability in a bladder cancer model. Microarray analysis is performed to profile probe expression changes in association with lung metastasis. Differentially expressed probes are applied to probe expression chromosome mapping analyses. These results are then compared to those obtained on the same xenograft model using Comparative Genomic Hybridization.*

comparison of locations of differentially expressed genes with that of locations of all spotted probes. In addition, to detect whether a particular chromosome had a high percentage of mapped genes that are significantly differentially expressed in association with tumor progression, we used a logistic regression in which the outcome was defined as whether or not a gene was significant. Because three chromosomes had no genes that were differentially expressed in the xenograft model, we added 0.5 to all cell counts. To discover if specific chromosomal regions contain differentially expressed genes at higher densities (physically concentrated; i.e., hot spots), the data were modeled as a nonstationary Poisson process. The model was implemented by applying a software typically used for counting process survival analysis [12,13]. Because the genes represented by probe sets in the microarray were unevenly distributed in the genome, it was necessary to account for the frequency of differentially expressed genes relative to the entire probe set distribution. Therefore, considering physical basepair distance as the metameter in survival analysis, derivation of the baseline survival curve using all probe sets for each chromosome comprised the control group. Evaluation of the hazard rate of differentially expressed genes in a similar fashion constituted the experimental group. In this setting, the test for proportional hazards assumption of the Cox model [14] detects whether the "hazard" of genes being differentially expressed (in the experimental group) is proportional to the hazard of genes being probed (in the control group). Thus, this approach analyzes whether there is any statistically meaningful unevenness in the distribution of differentially expressed genes in a chromosome once the distribution of probed genes has been taken into account. Because the test of proportional hazards simultaneously considers all the genes of a chromosome when comparing the significant ones with the baseline, there would be only one test performed; hence, no multiple-comparison issue arises. For figure generation and plots, differential expression data and annotations were exported from Affymetrix MAS 5.0 software and converted to text files. Custom scripts were written in Perl and R programming languages to render expression levels against chromosome positions. The National Center for Biotechnology Information data were used for chromosome lengths. The code for chromosome rendering was based on the Colored Chromosomes project [15].

Results

Gene Expression Mapping of Metastatic Phenotype in a Xenograft Model

We have previously described a bladder cancer metastasis model that represents a series of cell lines with progressively increased metastatic potential [7]. Genes whose altered expression was associated with metastatic progression were identified using high-density oligonucleotide microarrays containing ~22,500 probe sets. Of 18,513 evaluable probe sets, 164 were found to be significantly differentially expressed in association with lung metastasis. Sixteen of these probe sets did not have a chromosomal position assignment, which precluded their further analysis. An overview of gene expression mapping analysis and CGH is shown in Figure 1.

To detect whether a particular chromosome had a high percentage of mapped genes significantly differentially expressed with increasing lung metastasis, we used logistic regression in which the outcome was defined as whether or not a gene was significant. Overall, 0.8% of genes was found to be significant (Figure 2*A*). Chromosomes 2, 11, and 20 contained high percentages of differentially expressed genes (1.4%, 1.5%, and 2.0%, respectively; $P < .05$), in association with increasing lung metastasis, compared to the overall average. Results of counting process analysis (Table 1) confirmed that significant genes in chromosomes 2, 11, and 20 were distributed among the chromosomes in accordance with the baseline density of probed genes; hence, they had random physical distribution.

A gene (probe) is declared significantly differentially expressed once its expression is changed in all the three FL series in the same direction (i.e., up or down), compared to that of the T24T cell line. The significant proportion of genes was computed using both upregulated and downregulated genes, according to the above criteria. In particular, for chromosome 2, there were 2 downregulated and 15 upregulated significant genes; for chromosome, 11 there were 3 downregulated and 13 upregulated significant genes; and, for chromosome 20, there were 5 downregulated and 6 upregulated significant genes. However, *P* values for chromosomes 11 and 2 were dependent on the choice of the count added to cells (to alleviate zero counts). The significant proportions of genes in chromosomes 2 and 11 are closer to the overall average proportion across all the chromosomes, compared to the proportion of chromosome 20. When we added different values (such as 0.5) to cell counts to remedy zero cell counts, the proportions for chromosomes 2 and 11 fluctuated around the overall average, with *P* values fluctuating around the cutoff of 0.05. However, *P* values for the proportion of chromosome 20 remained stable. Hence, in a more conservative approach, we may declare only chromosome 20 to have a significantly higher percentage of significant genes.

A second analysis was performed to determine if any differentially expressed genes were clustered in specific regions in the chromosomes, thus constituting a hot spot—a chromosomal region with a high density of gene expression changes as a function of metastasis. One such region in chromosome 17q was identified, even though chromosome 17 did not contain the largest number of differentially expressed genes associated with metastatic progression (Table 1). On examination of the physical distribution of differentially expressed genes, eight were found to be unique genes (Table 2) clustered near the end of the long arm (17q21–17q25) (Figure 2*B*). Importantly, all but one of these genes, junction plakoglobin, were upregulated in association with metastatic progression. Repeating the hot-spot analysis using dChip provided a greater number of regions (Figure 3). Interestingly, only one area found in both techniques had a correlation with chromosomal changes by CGH.

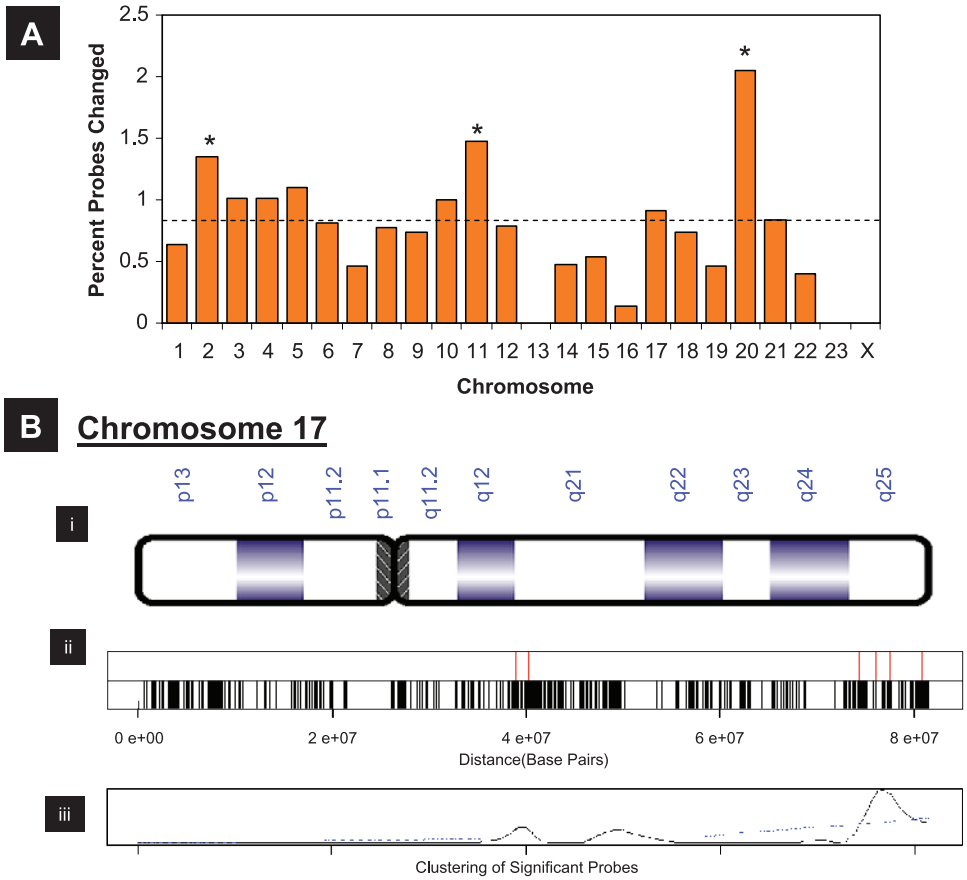

Figure 2. *Expression changes during metastatic progression and chromosomal position. (A) Percent differentially expressed probes during metastatic progression and chromosomal location. (–) Average percent change across all chromosomes.(*) Chromosomes harboring a statistically significantly higher number of differentially expressed probes with metastatic ability (see text for description). (B) Identification of the high-density area of gene expression changes in chromosome 17. (i) Chromosome ideogram. (ii) The distribution of all the array probes (black vertical lines) present in the chip, plus the distribution of the significantly upregulated or downregulated probes in T24T and FL3 tumor progression (red vertical lines). (iii) Graphical representation indicating where the significant probes are consistently concentrated, taking into account the baseline distribution of the probes. The y-axis represents scaled Schoenfeld residuals. The two curves are Lowess smoothers with different smoothing parameters fitted to the points. The dashed blue line smoothens more and points to the end of the chromosome, indicating that the general area of the hot spot is located at the end of the chromosome. The black solid line smoothens less and indicates a few peaks corresponding to hot spots, the largest of which is near the end of the q arm.*

Karyotypic Analysis of Metastatic Cell Lines

To determine whether chromosome-associated changes in gene expression were due to overt genetic alterations, we performed cytogenetic studies of the FL cell series and their parental cell line T24T. All of the cell lines had complex karyotypes with multiple structural and numerical abnormalities. The original cell line, T24T, contained between 73 and 80 chromosomes and between 3 and 6 marker chromosomes. Consistent abnormalities include: 73–80, (4n±) XXX, del(X)(q23), del(2)(p12),del(3)(p14), del(4)(q32), del(8)(p11.2), del(9)(q34), del(9)(p11.2), i(10)(q10), der(10)t(10;?)(q10;?), –15, der(17)t(17;8)(p11.1;q12), der(18)t(18;15)(p11.1;q12)(X2), +mar1, +mar2, +mar3, +mar4[cp5] (Figure 4). These were noted previously, with some refinements [8]. For example, the der(17)t(17:8)(p11.1;q12) identified now was formerly designated as a marker chromosome.

Compared to T24T, FL1 contained nearly all of the same numerical and structural abnormalities, with exceptions of a deletion of the short arm of chromosome 11 and additional copies of chromosome 22 (between three and five copies). The karyotypic designation of FL1 was therefore: 73–80, XXX, del(X)(q23), idem (T24T), del(11)(p11.1), +22, +22. The karyotype of FL2 was identical to that for FL1, but had an additional copy of chromosome 20 [73–80, XXX, del(X0(q23), idem (T24T), del(11)(p11.1), +20, +20, +22, +22]. The karyotype of FL3 was identical to that for FL2, except for a deletion of the short arm of chromosome 5 [73–80, XXX, del(X)(q23), idem (T24T), del(5)(p13), del(11)(p11.1), +20, +20, +22, +22].

For some chromosomes, there was a correlation between chromosomal alternations and gene expression. For example, chromosomes 11 and 20 had higher-than-average differentially expressed genes and chromosomal alterations associated with increasing lung metastasis (deletion of 11p in all FL series cell lines, but not in T24T; an additional copy of chromosome 20 was found in FL2 and FL3, but not in FL1 and T24T). For some chromosomes, there was no correlation between karyotypic alterations and differential gene expression. Additional copy numbers of chromosome 22 in the FL series *versus* T24T did not result in detectable differential gene expression from this chromosome. Although

Table 1. The Distribution of Differentially Expressed Probes and the Chromosomal Location of Hot Spots in the Xenograft Lung Metastasis Model of Human Bladder Cancer.

Chromosome	Probes	Differential Probes*	Hot-Spot *P* Value
1	1,892	12	.8223
2	1,258	17	.6203
3	993	10	.8542
4	694	7	.5973
5	815	9	.9480
6	1,110	9	.7215
7	855	4	.1303
8	644	5	.3682
9	675	5	.0301
10	703	7	.5889
11	1,087	16	.9123
12	1,010	8	.3152
13	342	0	NA
14	632	3	.3525
15	564	3	.2403
16	744	1	.1746
17	1,092	10	.0006
18	273	2	.0514
19	1,092	5	.5896
20	537	11	.4026
21	239	2	.2778
22	500	2	.1293
23	730	0	NA
Y	32	0	NA
Total	18,513	148	–

The hot-spot *P* value is calculated as described in Materials and Methods section. The significance level is $P < .01$.
*The number of significantly differentially expressed probes with consistent upregulation or downregulation among FL1, FL2, and FL3, when each compared to T24T, was identified by local pooled error test with $\alpha < .05$ [41].

chromosome 17 harbored a hot spot and chromosome 2 displayed a higher-than-average proportion of differentially expressed genes, both chromosomes did not have additional genetic changes in association with increasing lung metastasis.

Evaluation of Metastatic Cell Lines by CGH

CGH analyses demonstrated complex DNA changes in all cell lines (Table 3). The T24T cell line had a total of seven gains and four losses, and no HLA. FL1–FL3 cell lines had changes similar to those of T24T. However, these metastatic lesions were characterized by a progressive accumulation of additional gains and HLAs in specific chromosomal regions (Table 4). These cumulative genetic changes included gains at 1q, 3q, 7q, 8q, 9q, 12q, 17q, 18q, and HLA of 20q. There was no change in the loss profile of DNA copy number as a function of progression (Table 4).

DNA gains at 17q were detected in all metastatic FL1–FL3 cells. These gains were associated with our findings of nine overexpressed probes that mapped to 17q. The minimal common overlapping region using CGH was at 17q12-qter. The nine overexpressed genes mapped within this exact area. Although gains of chromosome 20 were seen in the nonmetastatic cell line (T24T), HLAs of 20q were characteristic of all metastatic cells. Microarray analysis indicated an overexpression of four probes that mapped to 20q. These constituted two different genes, transglutaminase 2 (*TGM2*) and BCL2-like 1 (*BCL2L1*). Although our results may be indicative of the role of increased DNA copy numbers in the overexpression of these genes, other upstream regulatory mechanisms cannot be ruled out.

Discussion

Bladder cancer can present either as a superficial lesion or as a muscularis propria–invasive tumor. Approximately half of the patients who present with invasive disease develop metastases despite initial treatment. Because genetic alterations and changes in gene expression are primary determinants controlling neoplastic transformation and progression, expression profiling using cDNA microarrays has been used in human cancers to aid in delineating the mechanisms involved in tumor progression [3]. In bladder cancer, gene expression profiles that are associated with tumor progression have been defined recently [2,16–18], but no studies have been conducted to compare gene expression changes with occurrences of specific genomic abnormalities. This is important because it is a common assumption that chromosomal rearrangements and imbalances seen in cancer cells affect gene expression, resulting in altered cellular phenotypes. The correlation between genetic alternations and gene expression changes has been studied in human breast tumors, head and neck squamous cell carcinomas, and gastric adenocarcinomas [6,19,20].

In this study, we used oligonucleotide microarrays to profile and evaluate gene expression changes in a model of

Table 2. Ten Probes Representing Eight Genes in the Long Arm of Chromosome 17 with Differential Expression in the Xenograft Lung Metastasis Model of Human Bladder Cancer.

Probe Set*	Gene Name	Map Location	Fold Difference†	*P* value‡
204990_s_at	Integrin, beta 4	17q11-qter	3.3	<.0001
201508_at	Insulin-like growth factor binding protein 4	17q12–q21.1	40.7	<.0001
201474_s_at	Integrin, alpha 3	17q21.33	2.5	<.0001
201015_s_at	Junction plakoglobin	17q21	−3.6	<.0001
200923_at	Lectin, galactoside-binding, soluble, 3 binding protein	17q25	5.8	<.0001
208657_s_at	MLL septin-like fusion	17q25	3.1	<.0001
203167_at	Tissue inhibitor of metalloproteinase 2	17q25	2.9	<.0001
41220_at	MLL septin-like fusion	17q25	2.1	<.0001
202855_s_at	Solute carrier family 16 (monocarboxylic acid transporters), member 3	17q25	1.8	<.0001
202856_s_at	Solute carrier family 16 (monocarboxylic acid transporters), member 3	17q25	1.7	<.0001

*Affymetrix (www.Affymetrix.com).
†Fold change comparing FL3 with T24T (FL3/T24T).
‡*P* value for local pooled error test (for testing the null hypothesis of equal mean gene expression across T24T, FL1, FL2, and FL3).

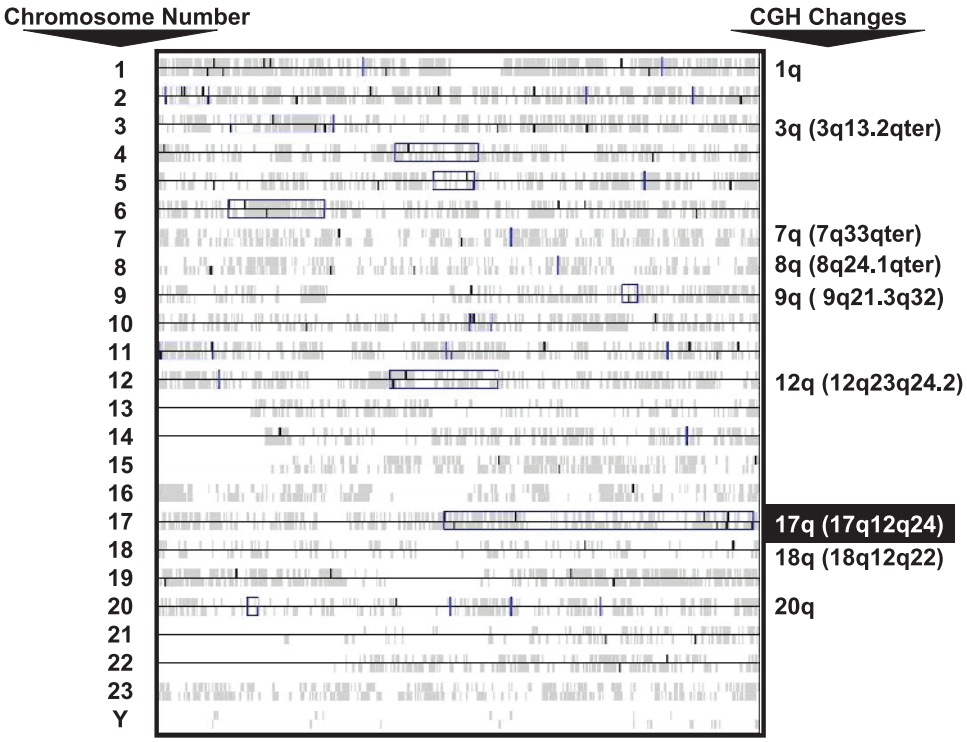

Figure 3. *dChip chromosomal clustering analysis of probe expression changes as a function of metastatic competence in the xenograft model. The program dChip was used to analyze the xenograft model as a function of metastasis. Hot spots were detected at the P = .02 level of significance (http://www.dchip.org), as outlined in blue boxes. Gray lines indicate probe positions, and black lines indicate probes that were differentially expressed. Observed CGH changes are presented (Table 4). Changes highlighted in black indicate cases wherein the dChip hot spot is in a location similar to that of the gains found in CGH.*

metastatic progression. We identified chromosomes 2, 11, and 20 as containing a higher percentage of differentially expressed genes as a function of metastatic progression. Interestingly, chromosome 20q also showed HLA on CGH analysis. Amplification of chromosome 20q has been associated with various cancers, including carcinomas of the

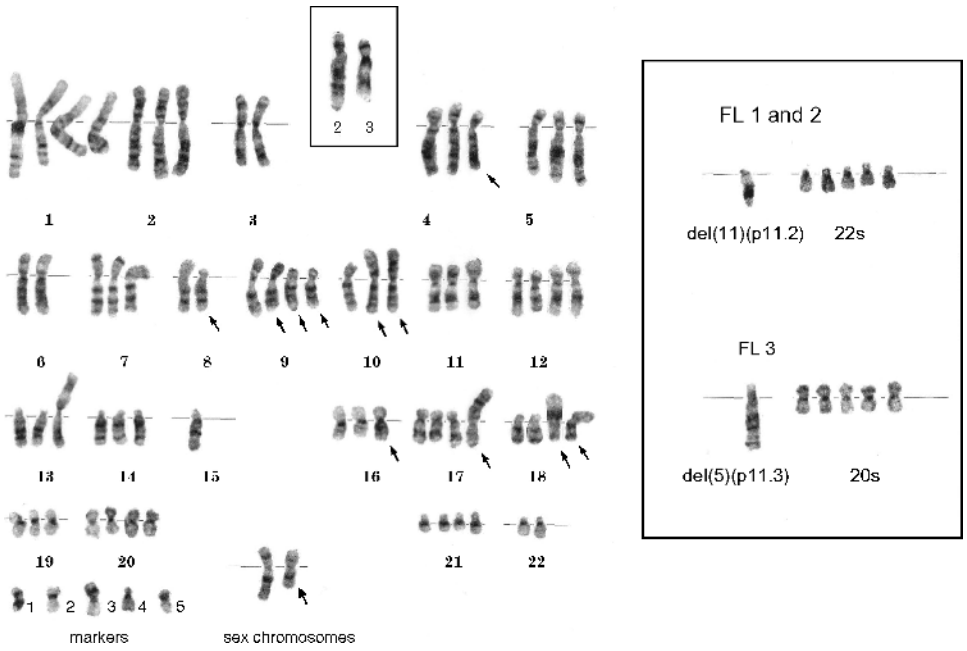

Figure 4. *Karyotypes of T24T and FL series. Cell lines were incubated for 5 to 7 days. When adequate mitotic cells had been observed, the coverslips were treated with Colcemid overnight and harvested* in situ*, using standard methods. The coverslips were stained using a standard trypsin–Giemsa banding method. Five karyotypes were produced from each cell line. A representative karyotype of T24T (composite) is shown. T24T has between 73 and 80 chromosomes, with several consistent structural abnormalities (arrows). The small insert illustrates the deleted chromosomes 2 and 3 that are not seen in this particular cell but are seen in several others. Larger inset reveals partial karyotypes of the FL1, FL2, and FL3 series, illustrating consistent changes seen compared to those of T24T.*

Table 3. CGH Changes Observed in the Xenograft Lung Metastasis Model of Human Bladder Cancer.

Cell Line	Gains	Losses
T24T	1p31.3pter, 2q34q36, 5p13pter, 5q31.3q34, 10q22.1q25.3, 11p12p15.1, 20	4q31.1qter, 8p, 9p13pter, 10p
FL1	1, 2q31q32.1, 3q13.2qter, 5p, 7q33qter, 9q21.3q32, 10q22.1qter, 11p, 12q14qter, 14q24.1q31, 17q12qter, 18q12q22, 20 (20q), 21q	4q31.1qter, 8p, 9p21pter, 10p, 19p
FL2	1, 3q13.2qter, 5p, 5q23.2qter, 7q33qter, 9q21.3q32, 10q22.1qter, 11p, 12q14qter, 17q12qter, 18q12q22, 20 (20q), 21q	4q31.1qter, 8p, 9p21pter, 10p
FL3	1, 3q13.2qter, 5p, 5q23.2qter, 7q33qter, 8q24.1qter, 9q21.3q32, 10q22.1qter, 11p, 12q14qter, 17q12qter, 18q12q22, 20 (20q)	4q31.1qter, 8p, 9p21pter, 10p

breast, colon, ovary, and pancreas [21–24]. In addition, a gain of chromosome 20q has been known to be associated with genome instability and tumor progression in bladder cancer [25,26]. Specific genes located in chromosome 20 that are related to bladder cancer progression include *STK15* (located in 20q13), which is associated with chromosomal aneuploidy [27,28]. Increased expression of *STK15* has been associated with bladder tumor grade, invasion, and metastasis [29]. Recently, another gene in chromosome 20, *CDC91L1*, was identified as an oncogene in bladder cancer and was found to be amplified and overexpressed in about one third of bladder cancer cell lines and primary tumors [30]. Finally, the ubiquitin conjugase UbcH10, located in 20q13.1, has been reported as significantly overexpressed in several human tumors, including bladder tumors [31].

To determine if there are areas within chromosomes that harbor unusually high levels of gene expression changes associated with lung metastasis, we applied a novel bioinformatics approach for the discovery and mapping of such gene expression hot spots. We identified chromosome 17q as harboring such hot spots that are concentrated between 17q12 and 17q22 and near the end of the long arm. This is consistent with evidence that clusters of functionally related genes could be located in particular chromosome regions, and that the expression of genes within a particular location could be affected by common mechanisms [32,33]. CGH analysis indicated DNA amplification of this area. Although alterations in 17q did not appear to correlate with grade, stage, or vascular invasion [34], losses [35], or gains [5,36], changes at 17q have been reported in transitional cell carcinomas compared to normal bladder tissues. Clustering of amplified and overexpressed genes at 17q has also been reported in upper gastrointestinal carcinomas (this study) [6].

Table 4. CGH Changes Observed during Metastatic Progression in the Xenograft Lung Metastasis Model of Human Bladder Cancer.

Chromosomal Location*	T24T	FL1	FL2	FL3
1q		G	G	G
3q (3q13.2qter)		G	G	G
7q (7q33qter)		G	G	G
8q (8q24.1qter)				G
9q (9q21.3q32)		G	G	G
12q (12q23q24.2)		G	G	G
17q (17q12q24)		G	G	G
18q (18q12q22)		G	G	G
20q	G	HLA	HLA	HLA

G, gains; HLA, high-level amplification.
*Chromosome arms are shown; the minimal overlapping region is indicated in parentheses.

Gene expression changes, of course, are not always associated with structural chromosomal abnormalities. We did not find a correlation between chromosomal alterations and gene expression for chromosome 2 despite the fact that chromosome 2 had a higher-than-average proportion of significantly differentially expressed genes. A lack of correlation between chromosomal abnormalities and expression profiles has also been reported by other groups [37,38], suggesting that mechanisms other than structural alterations influence gene expression in large genomic regions. This coordinated regional silencing or activation of multiple genes could be controlled by identified epigenetic changes that have been described in tumors [39,40].

The dChip software has a discovery tool for hot spots in gene expression data based on a nonparametric procedure for testing the uniformity of frequency for all segment regions of differentially expressed genes. Even though this tool can sensitively identify many potential locations of hot spots, it also provides a number of false positives due to multiple comparisons in a larger number of all segments of differentially expressed genes. To mitigate this pitfall, we have developed an overall testing procedure based on survival curve analysis that can rigorously evaluate the existence of hot spots in each chromosome. These two algorithms attempt to answer similar questions in that they aim to detect high (or low) frequencies of differentially expressed genes compared with spotty probe locations. However, they differ in that: 1) the former (counting process analysis) uses a semiparametric estimation of survival curves to evaluate the overall significance of such survival curve differences, whereas the latter (dChip) utilizes a nonparametric procedure for testing the uniformity of frequency for all segment regions of differentially expressed genes; and, consequently, 2) the former performs hypothesis testing on each chromosome (24 comparisons), whereas the latter does so for all consecutive segments within a chromosome (up to a certain length; e.g., 20 probes) for such a comparison. Therefore, as noted on the dChip web site, even though the dChip approach can effectively draw attention to some hot-spot locations, its *P* values need to be more carefully considered in terms of such multiple comparisons. The use of the dChip algorithm will be quite useful to identify the exact locations of hot spots after overall test methods, such as ours, have found statistical significance in certain chromosomes; the threshold *P* value of dChip can then be adjusted to reduce noise or to increase detection sensitivity compared to the overall test results. In our current study, by combining dChip and our novel survival curve–based analyses,

we have reliably determined that the hot-spot location in 17q is associated with metastatic phenotype in human bladder cancer.

In summary, oligonucleotide microarray analysis was performed to determine differential gene expression in a bladder cancer cell line metastasis model. Using a novel approach of chromosomal expression mapping, we found that gene expression changes with tumor metastasis do not occur randomly, but instead occur both in specific chromosomes and in intrachromosomal locations. Because some gene expression changes among chromosomes do not appear to be associated with CGH alterations, our data suggest that the presence of regional chromosomal mechanisms that control gene expression may be responsible for at least some of the changes observed during tumor progression.

Acknowledgements

The authors thank Jay Fox and Yongde Bao (University of Virginia Array Core Facility) for their assistance with chip hybridization.

References

[1] Ramaswamy S, Ross KN, Lander ES, and Golub TR (2003). A molecular signature of metastasis in primary solid tumors. *Nat Genet* **33,** 49–54.

[2] Dyrskjot L, Thykjaer T, Kruhoffer M, Jensen JL, Marcussen N, Hamilton-Dutoit S, Wolf H, and Orntoft TF (2003). Identifying distinct classes of bladder carcinoma using microarrays. *Nat Genet* **33,** 90–96.

[3] Liotta L and Petricoin E (2000). Molecular profiling of human cancer. *Nat Rev Genet* **1,** 48–56.

[4] Su AI, Welsh JB, Sapinoso LM, Kern SG, Dimitrov P, Lapp H, Schultz PG, Powell SM, Moskaluk CA, Frierson HF Jr, et al. (2001). Molecular classification of human carcinomas by use of gene expression signatures. *Cancer Res* **61,** 7388–7393.

[5] El-Rifai W, Kamel D, Larramendy ML, Shoman S, Gad Y, Baithun S, El-Awady M, Eissa S, Khaled H, Soloneski S, et al. (2000). DNA copy number changes in *Schistosoma*-associated and non–*Schistosoma*-associated bladder cancer. *Am J Pathol* **156,** 871–878.

[6] Koon N, Zaika A, Moskaluk CA, Frierson HF, Knuutila S, Powell SM, and El-Rifai W (2004). Clustering of molecular alterations in gastroesophageal carcinomas. *Neoplasia* **6,** 143–149.

[7] Nicholson BE, Frierson HF, Conaway MR, Seraj JM, Harding MA, Hampton GM, and Theodorescu D (2004). Profiling the evolution of human metastatic bladder cancer. *Cancer Res* **64,** 7813–7821.

[8] Gildea JJ, Golden WL, Harding MA, and Theodorescu D (2000). Genetic and phenotypic changes associated with the acquisition of tumorigenicity in human bladder cancer. *Genes Chromosomes Cancer* **27,** 252–263.

[9] Schneider BF, Lovell MA, and Golden WL (1998). Cytogenetic abnormalities in primary bronchopulmonary leiomyosarcoma of childhood. *Cancer Genet Cytogenet* **105,** 145–151.

[10] el-Rifai W, Larramendy ML, Bjorkqvist AM, Hemmer S, and Knuutila S (1997). Optimization of comparative genomic hybridization using fluorochrome conjugated to dCTP and dUTP nucleotides. *Lab Invest* **77,** 699–700.

[11] el-Rifai W, Sarlomo-Rikala M, Miettinen M, Knuutila S, and Andersson LC (1996). DNA copy number losses in chromosome 14: an early change in gastrointestinal stromal tumors. *Cancer Res* **56,** 3230–3233.

[12] Andersen P and Gill R (1982). Cox's regression model for counting processes: a large sample study. *Ann Stat* **10,** 1100–1120.

[13] Andersen PK, Borgan O, Gill RD, and Keiding N (1993). Statistical Models Based on Counting Processes Springer, New York.

[14] Grambsch P and Therneau T (1994). Proportional hazards tests and diagnostics based on weighted residuals. *Biometrika* **81,** 515–526.

[15] Bohringer S, Godde R, Bohringer D, Schulte T, and Epplen JT (2002). A software package for drawing ideograms automatically. *Online J Bioinformatics* **1,** 51–59.

[16] Modlich O, Prisack HB, Pitschke G, Ramp U, Ackermann R, Bojar H, Vogeli TA, and Grimm MO (2004). Identifying superficial, muscle-invasive, and metastasizing transitional cell carcinoma of the bladder: use of cDNA array analysis of gene expression profiles. *Clin Cancer Res* **10,** 3410–3421.

[17] Sanchez-Carbayo M, Socci ND, Lozano JJ, Li W, Charytonowicz E, Belbin TJ, Prystowsky MB, Ortiz AR, Childs G, and Cordon-Cardo C (2003). Gene discovery in bladder cancer progression using cDNA microarrays. *Am J Pathol* **163,** 505–516.

[18] Thykjaer T, Workman C, Kruhoffer M, Demtroder K, Wolf H, Andersen LD, Frederiksen CM, Knudsen S, and Orntoft TF (2001). Identification of gene expression patterns in superficial and invasive human bladder cancer. *Cancer Res* **61,** 2492–2499.

[19] Masayesva BG, Ha P, Garrett-Mayer E, Pilkington T, Mao R, Pevsner J, Speed T, Benoit N, Moon CS, Sidransky D, et al. (2004). Gene expression alterations over large chromosomal regions in cancers include multiple genes unrelated to malignant progression. *Proc Natl Acad Sci USA* **101,** 8715–8720.

[20] Pollack JR, Sorlie T, Perou CM, Rees CA, Jeffrey SS, Lonning PE, Tibshirani R, Botstein D, Borresen-Dale AL, and Brown PO (2002). Microarray analysis reveals a major direct role of DNA copy number alteration in the transcriptional program of human breast tumors. *Proc Natl Acad Sci USA* **99,** 12963–12968.

[21] Schlegel J, Stumm G, Scherthan H, Bocker T, Zirngibl H, Ruschoff J, and Hofstadter F (1995). Comparative genomic *in situ* hybridization of colon carcinomas with replication error. *Cancer Res* **55,** 6002–6005.

[22] Iwabuchi H, Sakamoto M, Sakunaga H, Ma YY, Carcangiu ML, Pinkel D, Yang-Feng TL, and Gray JW (1995). Genetic analysis of benign, low-grade, and high-grade ovarian tumors. *Cancer Res* **55,** 6172–6180.

[23] Solinas-Toldo S, Wallrapp C, Muller-Pillasch F, Bentz M, Gress T, and Lichter P (1996). Mapping of chromosomal imbalances in pancreatic carcinoma by comparative genomic hybridization. *Cancer Res* **56,** 3803–3807.

[24] Kallioniemi A, Kallioniemi OP, Piper J, Tanner M, Stokke T, Chen L, Smith HS, Pinkel D, Gray JW, and Waldman FM (1994). Detection and mapping of amplified DNA sequences in breast cancer by comparative genomic hybridization. *Proc Natl Acad Sci USA* **91,** 2156–2160.

[25] Prat E, Bernues M, Caballin MR, Egozcue J, Gelabert A, and Miro R (2001). Detection of chromosomal imbalances in papillary bladder tumors by comparative genomic hybridization. *Urology* **57,** 986–992.

[26] Bruch J, Wohr G, Hautmann R, Mattfeldt T, Bruderlein S, Moller P, Sauter S, Hameister H, Vogel W, and Paiss T (1998). Chromosomal changes during progression of transitional cell carcinoma of the bladder and delineation of the amplified interval on chromosome arm 8q. *Genes Chromosomes Cancer* **23,** 167–174.

[27] Sen S, Zhou H, and White RA (1997). A putative serine/threonine kinase encoding gene *BTAK* on chromosome 20q13 is amplified and overexpressed in human breast cancer cell lines. *Oncogene* **14,** 2195–2200.

[28] Zhou H, Kuang J, Zhong L, Kuo WL, Gray JW, Sahin A, Brinkley BR, and Sen S (1998). Tumour amplified kinase *STK15*/BTAK induces centrosome amplification, aneuploidy and transformation. *Nat Genet* **20,** 189–193.

[29] Sen S, Zhou H, Zhang RD, Yoon DS, Vakar-Lopez F, Ito S, Jiang F, Johnston D, Grossman HB, Ruifrok AC, et al. (2002). Amplification/overexpression of a mitotic kinase gene in human bladder cancer. *J Natl Cancer Inst* **94,** 1320–1329.

[30] Guo Z, Linn JF, Wu G, Anzick SL, Eisenberger CF, Halachmi S, Cohen Y, Fomenkov A, Hoque MO, Okami K, et al. (2004). *CDC91L1* (PIG-U) is a newly discovered oncogene in human bladder cancer. *Nat Med* **10,** 374–381.

[31] Wagner KW, Sapinoso LM, El-Rifai W, Frierson HF, Butz N, Mestan J, Hofmann F, Deveraux QL, and Hampton GM (2004). Overexpression, genomic amplification and therapeutic potential of inhibiting the UbcH10 ubiquitin conjugase in human carcinomas of diverse anatomic origin. *Oncogene* **23,** 6621–6629.

[32] Perou CM, Sorlie T, Eisen MB, van de Rijn M, Jeffrey SS, Rees CA, Pollack JR, Ross DT, Johnsen H, Akslen LA, et al. (2000). Molecular portraits of human breast tumours. *Nature* **406,** 747–752.

[33] Trowsdale J (1993). Genomic structure and function in the MHC. *Trends Genet* **9,** 117–122.

[34] Dalbagni G, Presti JC Jr, Reuter VE, Zhang ZF, Sarkis AS, Fair WR, and Cordon-Cardo C (1993). Molecular genetic alterations of chromo-

some 17 and p53 nuclear overexpression in human bladder cancer. *Diagn Mol Pathol* **2,** 4–13.

[35] Chaturvedi V, Li L, Hodges S, Johnston D, Ro JY, Logothetis C, von Eschenbach AC, Batsakis JG, and Czerniak B (1997). Superimposed histologic and genetic mapping of chromosome 17 alterations in human urinary bladder neoplasia. *Oncogene* **14,** 2059–2070.

[36] Richter J, Beffa L, Wagner U, Schraml P, Gasser TC, Moch H, Mihatsch MJ, and Sauter G (1998). Patterns of chromosomal imbalances in advanced urinary bladder cancer detected by comparative genomic hybridization. *Am J Pathol* **153,** 1615–1621.

[37] Phillips JL, Hayward SW, Wang Y, Vasselli J, Pavlovich C, Padilla-Nash H, Pezullo JR, Ghadimi BM, Grossfeld GD, Rivera A, et al. (2001). The consequences of chromosomal aneuploidy on gene expression profiles in a cell line model for prostate carcinogenesis. *Cancer Res* **61,** 8143–8149.

[38] Platzer P, Upender MB, Wilson K, Willis J, Lutterbaugh J, Nosrati A, Willson JK, Mack D, Ried T, and Markowitz S (2002). Silence of chromosomal amplifications in colon cancer. *Cancer Res* **62,** 1134–1138.

[39] Baylin SB and Herman JG (2000). DNA hypermethylation in tumorigenesis: epigenetics joins genetics. *Trends Genet* **16,** 168–174.

[40] Jones PA and Laird PW (1999). Cancer epigenetics comes of age. *Nat Genet* **21,** 163–167.

[41] Jain N, Thatte J, Braciale T, Ley K, O'Connell M, and Lee JK (2003). Local-pooled-error test for identifying differentially expressed genes with a small number of replicated microarrays. *Bioinformatics* **19,** 1945–1951.