Statistical Dissection of Genetic Pathways Involved in Prostate Carcinogenesis

Franclim R. Ribeiro,[1,2] Chieu B. Diep,[2] Carmen Jerónimo,[1,3] Rui Henrique,[4,5] Carlos Lopes,[4,5] Mette Eknæs,[2] Ole Christian Lingjærde,[6] Ragnhild A. Lothe,[2,7] and Manuel R. Teixeira[1,5]*

[1]Department of Genetics, Portuguese Oncology Institute, Porto, Portugal
[2]Department of Genetics, Institute for Cancer Research, Norwegian Radium Hospital, Oslo, Norway
[3]Fernando Pessoa University, Porto, Portugal
[4]Department of Pathology, Portuguese Oncology Institute, Porto, Portugal
[5]Institute of Biomedical Sciences, University of Porto, Portugal
[6]Bioinformatics Group, Department of Informatics, University of Oslo, Norway
[7]Department of Molecular Biosciences, University of Oslo, Oslo, Norway

Molecular markers that could stratify prostate cancer patients according to risk of disease progression would allow a significant improvement in the management of this clinically heterogeneous disease. In the present study, we analyzed the genetic profile of a consecutive series of 51 clinically confined prostate carcinomas and 27 benign prostatic hyperplasias using comparative genomic hybridization (CGH). We then added our findings to the existing literature data in order to perform a meta-analysis on a total of 294 prostate cancers with detailed CGH and clinicopathological information, using multivariate statistical methods that included principal component, hierarchical clustering, time of occurrence, and regression analyses. Whereas several genomic imbalances were shared by organ-confined, locally invasive, and metastatic prostate cancers, 6q and 10q losses and 7q and 8q gains were significantly more frequent in patients with extra-prostatic disease. Regression analysis indicated that 8q gain and 13q loss were the best predictors of locally invasive disease, whereas 8q gain and 6q and 10q losses were associated with metastatic disease. We propose a genetic pathway of prostate carcinogenesis with two distinct initiating events, namely, 8p and 13q losses. These primary imbalances are then preferentially followed by 8q gain and 6q, 16q, and 18q losses, which in turn are followed by a set of late events that make recurrent and metastatic prostate cancers genetically more complex. We conclude that significant differences exist in the genetic profile of organ-confined, locally invasive, and advanced prostate cancer and that genetic features may carry prognostic information independently of Gleason grade.

INTRODUCTION

Prostate cancer continues to represent a significant challenge to the clinical community worldwide, being one of the most frequently diagnosed malignancies and the second cause of cancer death among Western men (Jemal et al., 2004). As the majority of prostate carcinomas produce no clinical symptoms until late stages of progression, when therapeutic options are limited, strong emphasis has been put on early detection strategies. Prostate-specific antigen (PSA) screening programs clearly have improved detection of disease foci at a stage at which curative therapy can be provided. However, the clinical significance of many of these early tumor foci remains unclear (Etzioni et al., 2002). Current histopathological parameters cannot predict on an individual basis which of these organ-confined cancers will progress into a life-threatening condition and which will remain indolent during a patient's expected lifetime. As an aggressive treatment approach may seriously affect quality of life (Stanford et al., 2000; Steineck et al., 2002), there is urgent need to identify prognostic markers that can discriminate which prostatic carcinomas actually require active clinical management.

The relevance of genomic alterations for neoplastic development and progression has been widely acknowledged, and many efforts have been undertaken to uncover the genetic basis of prostate cancer progression (see Karayi and Markham, 2004, for a recent review). Comparative genomic hybridization (CGH), in particular, is a broad-spectrum molecular cytogenetic technique (Kallioniemi et al., 1992) that has been proven useful for identi-

Supported by: Fundação para a Ciência e a Tecnologia (FCT); Grant number: POCTI/CBO/38853/2001; Norwegian Cancer Society; Grant number: RAL: A95068; FCT research fellowship (to F.R.R.); Grant number: SFRH/BD7067/2001.

*Correspondence to: Manuel R. Teixeira, Department of Genetics, Portuguese Oncology Institute, Rua Dr. António Bernardino de Almeida, 4200-072 Porto, Portugal.
E-mail: mteixeir@ipoporto.min-saude.pt

Received 21 June 2005; Accepted 20 September 2005
DOI 10.1002/gcc.20279
Published online 18 October 2005 in Wiley InterScience (www.interscience.wiley.com).

fying chromosomal regions recurrently affected by genomic imbalances in prostate cancer. Losses of chromosome arms 8p, 13q, 6q, 10q, 16q, and 18q and gains of chromosome 7, 8q, and Xq have been reported by several groups (Visakorpi et al., 1995; Cher et al., 1996; Alers et al., 2000, 2001; Fu et al., 2000; Mattfeldt et al., 2002; Wolter et al., 2002a, 2002b; Chu et al., 2003; Teixeira et al., 2004). Many of these series, however, were selected to represent specific stages or subgroups of the disease, making it difficult to compare findings and derive clinically relevant conclusions.

In the present study, we used CGH to analyze a prospective series of 51 fresh-frozen samples from patients with clinically confined prostate cancer that had consecutively undergone radical prostatectomy at our institution. Additionally, a series of 27 benign prostatic hyperplasia (BPH) samples was studied as a control group that might also display raised PSA levels. We then performed a meta-analysis that included a total data set of 418 cases with individual CGH data, of which 294 also had individual clinicopathological information. Through multivariate statistical analysis of this data set, we aimed to uncover significant differences in the genetic profile of organ-confined, locally invasive, and advanced prostate cancers in order to identify genetic features that might carry independent prognostic information.

One of the statistical methods applied to the metadata was time of occurrence (TO) analysis. The concept of statistically determining a temporal order of acquisition of genetic imbalances was put forward by Höglund et al. and has provided interesting clues regarding the contribution of specific genetic changes to the development and progression of many tumor types (Höglund et al., 2004a, 2004b, 2005). Existing TO analyses are based on cytogenetic data extracted from published karyotypic information, which is scarce for prostate tumors (Mitelman et al., 2005). To arrive at a genetic model of prostate carcinogenesis, we applied the TO principle and methodology with the modifications described by Diep et al. (2005) to the entire data set of copy number changes assessed by CGH.

MATERIAL AND METHODS

Prostate Carcinomas and Benign Hyperplasias

A collection of fresh-frozen tissue from patients undergoing radical prostatectomy at the Portuguese Oncology Institute–Porto has been created and maintained since 1999. For the purposes of this study, 54 consecutive prostatectomy samples (obtained between September 1999 and June 2000) from patients with clinically confined prostate cancer (T1c and T2) who had not received preoperative therapy were chosen. After surgical resection of the prostate and seminal vesicles, the removed gland was thickly sliced and visually inspected for tumor presence. Fragments from suspected cancer foci were sectioned and fresh-frozen at −80°C. Five-micron sections of the frozen fragments were cut, stained, and evaluated by a pathologist. The frozen-tissue block was then trimmed to maximize the yield of target cells (>70% tumor content) and, on average, twenty 12-μm-thick sections were obtained for DNA extraction, with every 5th section stained in order to ensure tumor representativeness. Three samples were dismissed because there was insufficient DNA for good-quality CGH analysis. Pathological staging of the remaining 51 samples (Table 1) followed the pTNM system (Hermanek et al., 1997), and grade was assessed according to the Gleason score (GS; Gleason and Mellinger, 1974). The median ages of patients with pT2 and pT3 tumors were 63 (range 46–73) and 64 (range 51–74), respectively, whereas median PSA levels for these stages were 7.5 (range 4.5–16.9) and 9.4 (range 3.71–17.45), respectively. Thirty randomly selected samples from patients with BPH who had undergone transurethral resection of the prostate were also analyzed. Twenty-three of these were fresh-frozen, whereas seven corresponded to paraffin-embedded tissue. Patients had received no treatment before resection. Sample blocks were evaluated by a pathologist to confirm diagnosis and ensure the absence of malignant neoplastic tissue, after which, on average, twenty 12-μm sections were obtained for DNA extraction. Three paraffin samples were discarded because there was insufficient DNA for good-quality CGH analysis. The median age of the BPH patients was 68 (range 54–79 years), and their median PSA level was 5.75 (0.9–12.7).

Comparative Genomic Hybridization

CGH analysis followed the procedure of Kallioniemi et al. (1992) with modifications previously described by Kraggerud et al. (2000) and Teixeira et al. (2004). For the scoring of CGH results, we recently adopted the use of dynamic standard reference intervals (D-SRIs). A D-SRI represents a "normal" ratio profile that encompasses the amount of variation detected in negative controls for each chromosome band. It has been shown to provide a more objective and sensitive scoring criterion when compared with fixed thresholds (Kirchhoff et al., 1998, 1999). The D-SRI used was

TABLE 1. Clinicopathological Data and Corresponding Genomic Findings on 51 Prostate Carcinomas

Age	PSA	pT Staging	Gleason score	CGH findings (standard reference intervals 99%)
63	6.40	2a	6 (3 + 3)	No changes
46	9.38	2a	6 (3 + 3)	No changes
51	4.63	2a	6 (3 + 3)	rev ish dim(5q23q31,8p21p22)
65	6.33	2a	7 (3 + 4)	rev ish enh(5p14pter,5q11q23,5q32q33)
68	12.70	2a	7 (3 + 4)	rev ish enh(8q),dim(2q23q24,8p12p23,10p11p12,10q22q25)
63	9.46	2a	8 (4 + 4)	rev ish dim(8p12pter)
64	12.20	2b	4 (2 + 2)	rev ish dim(6p21,8p12pter,13q14q31)
55	8.00	2b	6 (3 + 3)	No changes
60	10.30	2b	6 (3 + 3)	No changes
67	16.90	2b	6 (3 + 3)	No changes
69	7.79	2b	6 (3 + 3)	No changes
60	7.10	2b	6 (3 + 3)	No changes
58	4.99	2b	6 (3 + 3)	No changes
55	12.68	2b	6 (3 + 3)	No changes
47	4.80	2b	6 (3 + 3)	No changes
65	5.50	2b	6 (3 + 3)	No changes
70	7.10	2b	6 (3 + 3)	No changes
73	7.00	2b	6 (3 + 3)	No changes
64	4.25	2b	6 (3 + 3)	No changes
65	11.20	2b	7 (3 + 4)	No changes
63	6.76	2b	7 (3 + 4)	No changes
64	11.60	2b	7 (3 + 4)	rev ish dim(8p12pter)
62	4.80	2b	7 (3 + 4)	No changes
62	6.49	2b	7 (3 + 4)	No changes
70	10.83	2b	7 (3 + 4)	No changes
67	4.80	2b	7 (3 + 4)	No changes
67	9.97	2b	7 (3 + 4)	No changes
66	6.90	2b	7 (3 + 4)	rev ish dim(5q21q23,8p21p22)
60	12.00	2b	7 (3 + 4)	No changes
67	12.88	2b	7 (3 + 4)	rev ish dim(6q15q21,8p21p23,13q21q31)
54	15.40	3a	6 (3 + 3)	rev ish dim(5q31,8p21p22,16q23q24)
63	9.30	3a	6 (3 + 3)	rev ish dim(8p12p23)
66	7.92	3a	6 (3 + 3)	rev ish dim(8p12pter)
61	11.14	3a	6 (3 + 3)	No changes
64	10.00	3a	6 (3 + 3)	rev ish dim(8p21pter)
51	13.00	3a	6 (3 + 3)	rev ish dim(16q22qter)
60	17.45	3a	7 (3 + 4)	rev ish dim(8p21p22,10q22q23)
69	6.80	3a	7 (3 + 4)	rev ish dim(8p12p22)
62	7.60	3a	7 (3 + 4)	No changes
66	8.50	3a	7 (3 + 4)	No changes
53	10.00	3a	7 (3 + 4)	rev ish enh(18p11),dim(8p22pter,13q14q22)
69	9.60	3a	7 (3 + 4)	rev ish dim(8p22pter,17p13)
70	9.60	3a	7 (3 + 4)	rev ish dim(6q15q22)
59	7.20	3a	7 (3 + 4)	No changes
62	3.71	3a	7 (3 + 4)	rev ish dim(8p12p22,17p12pter)
66	17.00	3a	7 (4 + 3)	No changes
74	11.50	3a	8 (3 + 5)	rev ish dim(8p11pter,16q22q24)
68	nd	3a	8 (3 + 5)	No changes
66	11.60	3a	9 (4 + 5)	rev ish enh(3q23q26,7p13p21,7q21q32,8q21q24),dim(16q22qter)
68	9.30	3b	7 (3 + 4)	rev ish dim(8p12p22,10q22qter,13q14q21,16q23q24)
65	7.20	3b	8 (3 + 5)	rev ish enh(8q21q24),dim(8p22)

Preoperative clinical stage was T_{1c} or T_{2a} (N_0M_0) for all cases (nd: not determined).

generated with data from 10 normal versus normal hybridizations (totaling 110 cells). This interval was automatically scaled onto each sample profile, and aberrations were scored whenever the case profile and the standard reference profile at 99% confidence did not overlap. Description of the CGH copy number changes (CNC) followed the guidelines suggested in the ISCN (1995).

Meta-analysis of Prostate Cancer CGH Data

In January 2005, an extensive literature search was performed on PubMed (http://www.ncbi.nlm.

nih.gov/PubMed/) with the purpose of identifying all publications dealing with chromosome CGH analysis of clinical prostate cancer samples. Twenty-seven articles were identified using different keywords and cross-checking all references to decrease the possibility of missing relevant publications. From these 27 publications, however, only 15 provided either individual CGH data for their samples or graphical diagrams that would allow us to mine this information (Visakorpi et al., 1995; Nupponen et al., 1998; Alers et al., 2000, 2001; Fu et al., 2000; Alers et al., 2001; El Gedaily et al., 2001; Rokman et al., 2001; Kasahara et al., 2002; Schulz et al., 2002; Wolter et al., 2002a, 2002b; Chu et al., 2003; Van Dekken et al., 2003; Matsuda et al., 2004; Teixeira et al., 2004). Furthermore, as one of the clinical endpoints of this work was to genetically distinguish pathologically organ-confined tumors from locally invasive or metastatic cancers, only samples that could without a doubt be attributed to one of these categories were selected. T1 and pT2 samples were placed in the group of gland-confined cancers, pT3-4N0M0 samples were grouped as locally invasive cancers, and metastases and recurrences were grouped as advanced cancers. When Gleason scores were available, samples were categorized into three groups ($GS \leq 6$, $GS = 7$, and $GS \geq 8$). After pooling our cases with those reported in the literature, individual CGH data were available for a total of 418 cases (274 with CNC), with both individual clinicopathological and CGH data available for 294 of these (176 with CNC). The Gleason score was available for 201 samples (102 with CNC). The data were then coded by chromosome arm as having the absence (0) or presence (1) of a given imbalance, resulting in a spreadsheet with 82 data points for each sample (41 gains, 41 losses; the p arms of the acrocentric chromosomes as well as the Y chromosome were excluded from this analysis). Subsequent statistical tests were performed on the data matrices thus created. The chi-square test, chi-square test for trend, and Fisher's exact test were applied according to the categorization of the variables. The Kruskal–Wallis and Mann–Whitney nonparametric tests were used to assess the relationship between the total number of genomic imbalances and tumor grade and stage, according to the number of groups compared. A P value smaller than 0.05 (two-sided) was considered an indication of statistical significance, except when chromosome arms affected in more than 10% of the cases were compared with tumor grade or stage, in which case a Bonferroni correction was introduced to account for multiple testing. All analyses were performed using SPSS, version 11.0 (SPSS, Chicago, IL).

Time of Occurrence Analysis

Time of occurrence analysis was performed essentially as described by Höglund et al. (2001) with the modifications suggested by Diep et al. (2005). Briefly, if it is assumed that imbalances are early if predominantly seen in samples with few other copy number changes and are late if predominantly observed in samples with a higher number of changes, a temporal pattern of acquisition of these aberrations during tumor progression can be inferred. This analysis was performed in the subset of 274 samples with CNC for which CGH information on individual cases was available. All imbalances present in more than 2.5% of the available samples were included. A distribution of the number of imbalances per tumor (NIPT) was then generated, and the mode of this distribution was calculated for each imbalance. These values were then used as an indication of the order in which these aberrations occurred during prostate carcinogenesis, referred to as time of occurrence (TO). To evaluate if the observed TO values were significantly different from what would be expected from random events, a bootstrapping simulation procedure was performed. Imbalances were considered significantly early or late whenever the observed TO value was lower or higher than the 2.5% and 97.5% limits of the confidence interval, respectively, obtained in the simulation. Resampling was performed in Java (Sun Microsystems, Santa Clara, CA).

Multivariate Statistics

Several multivariate analyses were used to simplify data interpretation and to assess the independent contribution of each individual variable. Hierarchical clustering analysis (HCA), through the use of standard statistical algorithms, arranges genetic aberrations and tumor samples in groups based on their similarity, with the results visualized in the form of a dendogram (Aldenderfer and Blashfield, 1984). The average-linkage method with Pearson's correlation similarity measure was used. Principal component analysis (PCA) looks for underlying patterns in large data sets by reducing the overall number of variables to a limited number of components that best explain the variance within the data set (Affifi and Azen, 1979). To introduce a temporal order in PCA, the variable TO was also entered into the model (Höglund et al., 2002). These analyses were performed in

J-Express Pro 2.5 (Dysvik and Jonassen, 2001) using the subset of 274 abnormal cases. Finally, to evaluate the relative contribution of genetic and histologic variables in the assessment of pathological status, multivariate logistic regression was performed in SPSS. The three groups of samples (organ confined, locally invasive, and recurrent/metastatic) were compared on a two-by-two basis in the model, and several independent variables (chromosomal aberrations seen in more than 10% of the cases, presence or absence of genomic imbalances, degree of genetic complexity, and categorized Gleason score) were then used in a forward conditional setting to establish the best predictors of the dependent variables. Evaluation of genetic features as prognostic markers was performed on a total of 294 samples (176 with CNC).

Genetic Pathway of Prostate Carcinogenesis

Taking into consideration the results from the different statistical methods, we constructed a tentative genetic pathway for prostate cancer development and progression. Only genetic alterations present in at least 10% of the cases with CNC were included in the model. These were categorized according to the PCA into early, intermediate, or late events. The relationship between different variables (as indicated by HCA results) was also taken into account to generate the final model.

RESULTS

Original Genomic Findings

Copy number changes were detected in 23 of 51 prostate carcinomas (45%). Losses were seen in 22 tumors (43%), whereas only five cases (10%) showed DNA copy number gains. Among the abnormal cases, the average number of genomic imbalances was 2.2 (1.8 for losses and 0.4 for gains). When the relationship between genetic and histopathological variables was assessed, the number of cases presenting CNC along the three GS categories increased significantly (32%, 50%, and 80%, $P = 0.044$). An increased frequency of genomic gains was also seen in the subset of poorly differentiated tumors ($P = 0.008$), most often 8q gain ($P = 0.006$). The number of cases presenting DNA copy number changes was significantly higher in the subset of tumors with extra-capsular invasion ($P = 0.002$), as was the frequency of cases with losses ($P = 0.001$), especially of 8p ($P = 0.02$) and 16q ($P = 0.009$). Detailed genomic findings of the 51 prostate carcinomas and the respective clinicopathological data are presented in Table 1.

Meta-analysis of CGH Data and Clinicopathological Correlations

A graphical comparison of the frequencies of individual chromosomal aberrations among organ-confined ($n = 170$), locally invasive ($n = 69$), and metastatic cancers ($n = 55$) is depicted in Figure 1, and the relevant histopathological correlations are presented in Table 2. The percentage of cases displaying DNA copy number changes increased significantly along these three groups of patients (45%, 72%, and 89%, respectively; $P < 0.001$). The average number of aberrations in the abnormal cases was 3.0 (2.0 losses and 1.0 gains) for organ-confined, 5.3 (3.1 losses and 2.2 gains) for locally invasive, and 10.7 (5.7 losses and 5.0 gains) for metastatic/recurrent cancers ($P < 0.001$). The frequency of gains at chromosomal arms 7q and 8q and losses at 6q and 10q increased significantly along these groups of patients ($P < 0.003$). Among the three GS categories, the average number of CNC was 2.9, 3.1, and 5.3, respectively ($P < 0.036$). As expected, an association was also found between GS and pathological stage ($P < 0.001$). Of note, Xq gain was mostly observed in recurrences and metastases after androgen-deprivation therapy.

To determine which of the individual genetic variables would retain significance in multivariate analysis, a stepwise binary logistic regression was performed using the three disease stages as dependent variables (Table 3). When individual genetic variables were introduced together with GS, 8q gain and 13q loss would differentiate best between organ-confined and locally invasive cancers. Loss of 6q, on one side, and 8q gain and 10q loss, on the other, were associated with recurrent/metastatic disease as compared to organ-confined and locally invasive disease, respectively. Genomic complexity, coded as tumors displaying more than two genomic changes, would also be selected in the model to differentiate organ-confined or locally invasive disease from metastatic cancers (together with 7q gain and 10q loss).

Genetic Pathway of Prostate Cancer Progression

A total of 14 imbalances were recurrently found in more than 10% of the 274 abnormal samples for which individual CGH data were available. This threshold was chosen in order to remove the possible background generated by less frequent imbalances. Temporal analysis of this data set revealed that losses at 8p, 13q and gains at 16q and 17q are significantly early events, followed by 6q loss and 8q gain. Significantly late genomic changes

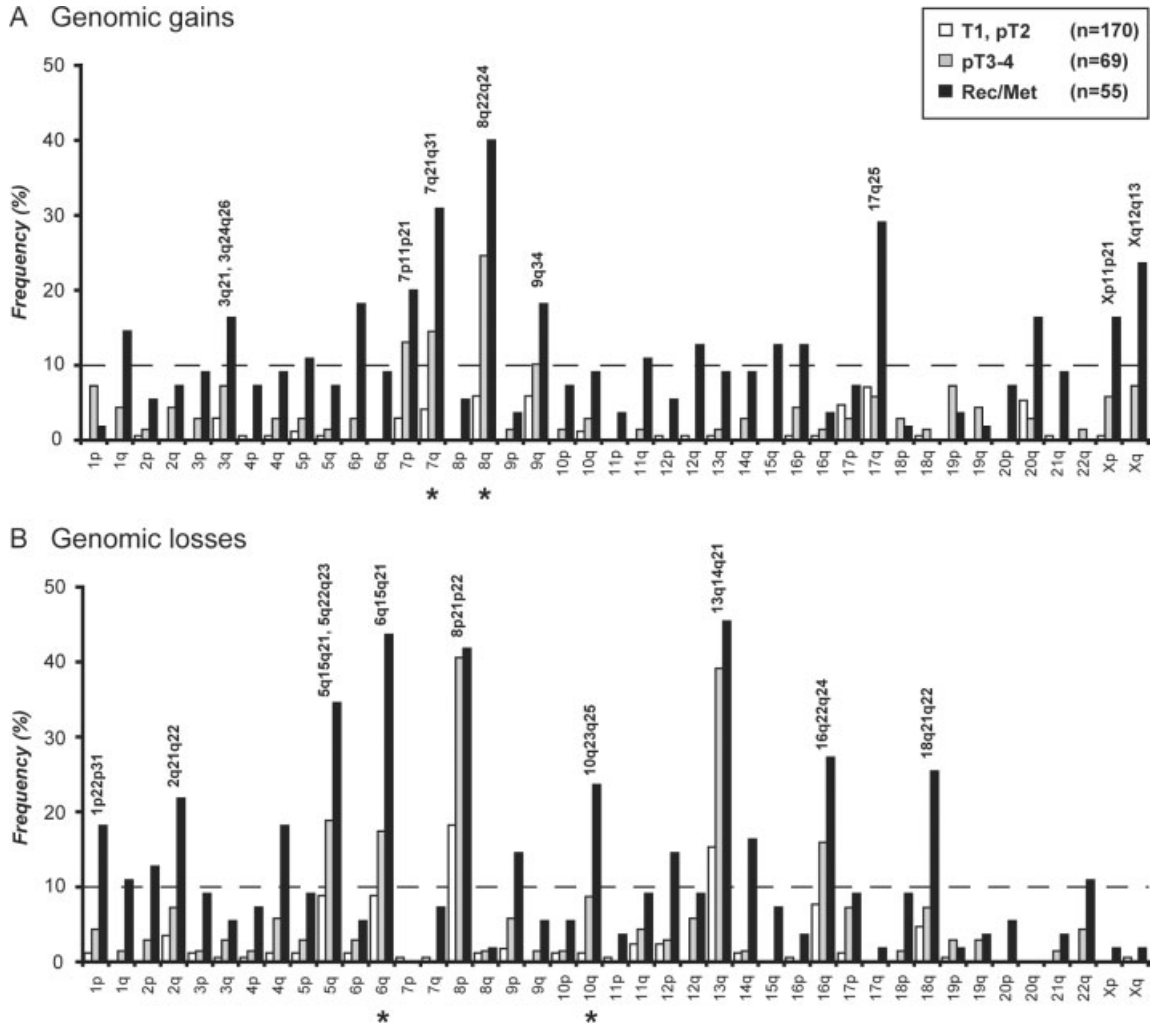

Figure 1. Frequency of copy number changes in prostate cancer (meta-analysis of 294 cases from 15 studies and the present study). Gains (A) and losses (B) are depicted, classified by chromosome arm. For each aberration occurring in more than 10% of cases, the minimal regions of overlap are indicated. Asterisks indicate statistically significant differences among the three groups of samples (after Bonferroni correction for multiple testing).

included gains at 7p and 7q. PCA on the same data set (using the variable TO) indicated that 8p and 13q losses are independent initiating events in prostate carcinogenesis, followed by a group of four intermediate alterations (8q gain and 6q, 16q, and 18q losses) and then by a tighter cluster of late events that included all the other frequently gained regions and also losses at 2q, 5q, and 10q (Fig. 2A). Hierarchical clustering of the same data set also showed a separation between 8p and 13q losses, further suggesting that 16q loss is frequently associated with 8p loss, whereas 6q and 18q losses are more closely associated with 13q loss (Fig. 2B).

On the basis of the findings detailed above, a tentative genetic pathway of prostate cancer progression is proposed in Figure 3. Ninety-three percent of the prostate carcinomas with CNC in the metadata ($n = 274$) displayed at least one of the most common genomic imbalances in this model, of which 73% displayed at least one early, 50% presented at least one early and one intermediate, and 29% showed a combination of early, intermediate, and late genetic events. Of the cases with 8p loss as the only early event (82/274), 38% acquired either 8q gain or 16q loss, whereas 27% acquired 6q or 18q losses instead ($P = 0.181$). Of the cases with 13q loss as the only early event (73/274), 48% acquired 6q or 18q losses, whereas only 26% harbored 8q gain or 16q loss ($P = 0.0098$). The acquisition of the intermediate events 6q loss or 18q loss was significantly associated with the early event 13q loss when compared to 8p loss ($P = 0.0077$).

TABLE 2. Relationship Between Genomic Imbalances and Histopathological Data

	Gleason score				Pathological status			
	4–6	7	8–10	*P* value	Organ confined	Locally invasive	Recurrent/metastatic	*P* value
	$n = 50$	$n = 40$	$n = 12$		$n = 77$	$n = 50$	$n = 49$	
Number of CNC[a]	2.9	3.1	5.3	0.036	3.0	5.3	10.7	<0.001
Number of gains[a]	1.1	0.9	2.6	n.s.	1.0	2.2	5.0	<0.001
Number of losses[a]	1.8	2.2	2.7	n.s.	2.0	3.1	5.7	<0.001
Cases with 2q loss	2%	10%	17%	n.s.	8%	10%	24%	n.s.
Cases with 5q loss	16%	25%	25%	n.s.	20%	26%	39%	n.s.
Cases with 6q loss	20%	15%	33%	n.s.	20%	24%	49%	0.001
Cases with 8p loss	32%	58%	83%	<0.001	40%	56%	47%	n.s.
Cases with 10q loss	0%	8%	8%	n.s.	3%	12%	27%	<0.001
Cases with 13q loss	36%	35%	33%	n.s.	34%	54%	51%	n.s.
Cases with 16q loss	20%	15%	25%	n.s.	17%	22%	31%	n.s.
Cases with 18q loss	6%	10%	17%	n.s.	10%	10%	29%	n.s.
Cases with 3q gain	6%	13%	17%	n.s.	7%	10%	18%	n.s.
Cases with 7q gain	4%	8%	25%	n.s.	7%	18%	22%	n.s.
Cases with 7q gain	4%	10%	33%	n.s.	9%	20%	35%	<0.001
Cases with 8q gain	10%	13%	58%	0.001	13%	34%	45%	<0.001
Cases with 9q gain	14%	8%	25%	n.s.	13%	14%	20%	n.s.
Cases with 17q gain	20%	3%	17%	n.s.	16%	8%	33%	n.s.

[a]Mean values for each category are displayed. CNC: copy number changes; n.s.: not significant. Only *P* values that retain significance after correction for multiple testing are indicated.

TABLE 3. Regression Analysis Using Disease Stage as Dependent Variable and Genetic Features as Independent Variables

Disease status[a]	Genetic event	*P* value	OR	95% CI for OR
Organ confined versus locally invasive	Gain 8q	0.001	3.651	1.724–7.734
	Loss 13q	0.024	2.375	1.121–5.030
Locally invasive versus recurrent/metastatic	Loss 6q	0.013	2.808	1.244–6.339
Organ confined versus recurrent/metastatic	Gain 8q	0.002	3.759	1.642–8.603
	Loss 10q	0.004	3.353	1.465–7.673

[a]Organ-confined cancers: $n = 77$ abnormal cases; locally invasive cancers: $n = 50$; recurrent/metastatic cancers: $n = 49$. OR: odds ratio; CI: confidence interval. The variable Gleason score, available for most organ-confined and locally invasive cancers, was entered into the model together with the genetic variables.

DISCUSSION

We report a genomewide survey of DNA copy number changes in a consecutive series of clinically localized prostate cancers and randomly selected benign prostatic hyperplasias. The absence of genomic imbalances in BPH presents further evidence that this clinical entity is not implicated in prostate carcinogenesis. Interestingly, more than 70% of the pT2 tumors also did not display chromosome-level imbalances. Because these tumors were carefully dissected to maximize the yield of neoplastic tissue, this finding could indicate that submicroscopic genetic or epigenetic changes are responsible for these morphologically confirmed cancers or that this subset of prostate carcinomas harbor only balanced cytogenetic changes that cannot be detected by CGH. Our previous chromosome banding analyses of prostate cancer indicate that balanced karyotypic abnormalities are rare (Teixeira et al., 2000, 2004), thus arguing against that latter hypothesis. In addition, we have previously demonstrated promoter hypermethylation of *GSTP1* in these very same tumors (Jeronimo et al., 2001), thus indicating that epigenetic mechanisms may be operative in these prostate carcinomas without chromosome-level changes. These findings are in accordance with other reports in the literature, in which the percentage of cases without DNA copy number changes ranged from 40% to 50% in clinically confined tumors (Fu et al., 2000; Alers et al., 2001; Chu et al., 2003) to 70% in incidental carcinomas (Wolter et al., 2002a). Further genomic investigations with higher-resolution methodologies, like array-based CGH, will be necessary to characterize genetically this subset of prostate cancers.

The correlations between genetic findings and clinicopathological variables using the meta-analysis data allowed us to derive two main conclusions. First, genetically complex prostate cancers are associated with higher pathological or clinical stag-

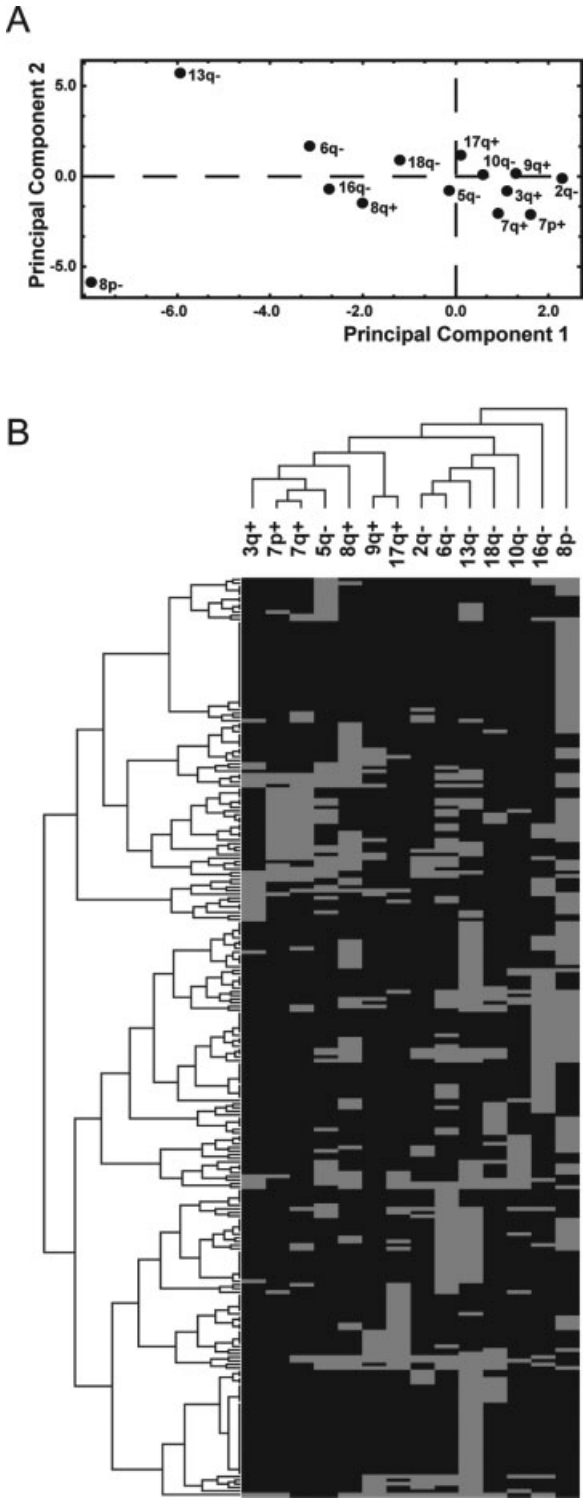

Figure 2. A. Principal component analysis (PCA) using TO values as an observation. Only genetic alterations seen in more than 10% of the cases ($n = 274$) are depicted. These are ordered along principal component 1, which corresponds to the TO variable; B. Hierarchical clustering of the imbalances seen in more than 10% of the prostate cancers (top dendogram). The left dendogram shows clustering of prostate cancers based on the genomic imbalances.

ing. Interestingly, the relative proportion of gains and losses gradually shifted along the different disease stages, with organ-confined tumors presenting twice the amount of losses per sample and metastatic cancers showing almost the same proportion of gains and losses per sample. Second, the specific genomic imbalances 8q gain and 13q loss are associated with extraprostatic disease independently of Gleason grade, which indicates that genetic features may add significant prognostic information to standard histopathological analysis of prostate cancer specimens. Prognostic significance of DNA copy number changes has previously been suggested by other investigators (Alers et al., 2000; Fu et al., 2000; Steiner et al., 2002; Van Dekken et al., 2003), but these studies were either retrospective or used biochemical progression as the clinical end point. Prospective investigations with sufficiently long follow-up to use survival as the clinical end point will be necessary to determine the clinical usefulness of genetic prognostic markers for prostate cancer patients.

To infer a temporal order of the genetic events relevant for prostate carcinogenesis, we have combined the results from the multivariate analysis PCA (using the TO as a variable) and HCA on the metadata. On the basis of the information thus obtained, we propose that genomic losses at 8p and 13q are distinct initiating events in prostate carcinogenesis, something that is also supported by the fact that these are the only genomic imbalances recurrently found as the sole abnormality in prostate carcinomas. The loss of 8p is preferentially followed by 16q loss and 8q gain, whereas 13q loss is preferentially followed by 6q and 18q losses (Fig. 3). As tumors progress, the two pathways converge by the acquisition of several secondary DNA copy number changes, making recurrent and metastatic cancer genetically more complex and heterogeneous. Interestingly, high-grade prostatic intraepithelial neoplasia (PIN), the most likely precursor of prostate cancer, has been shown to frequently harbor 8p deletions by loss of heterozygosity and FISH studies (see Foster et al., 2000, for a thorough review). Conversely, loss of 13q is rarely found in PIN and has in fact been associated with advanced stages of prostate cancer (Dong et al., 2000). Because we have shown in the present study that a significant proportion of prostate cancers present a 13q deletion as an early event instead of an 8p deletion and that there is no convincing evidence for other premalignant lesions in prostate carcinogenesis (Doll et al., 1999; Foster et al., 2000; Bettendorf et al., 2005), it is tempting to pro-

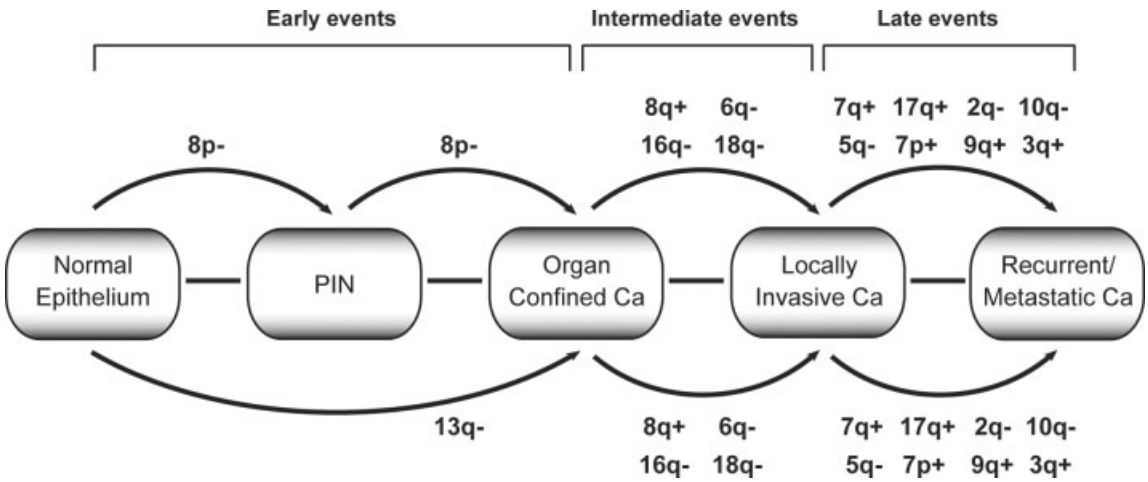

Figure 3. Genetic model of prostate cancer progression based on genomic imbalances detected by comparative genomic hybridization. DNA copy number changes detected in more than 10% of the abnormal prostate cancer samples were categorized as early, intermediate, or late events according to time of occurrence and principal component analyses. Two preferential pathways of genetic progression are proposed, one starting with 8p loss and the other starting with 13q loss. See text for further information.

pose that a distinct pathogenetic pathway in prostate carcinogenesis starts with a 13q deletion and progresses to local invasion and metastasis without going through a stage of intraepithelial proliferation.

Although we have shown that preferential genetic pathways do exist in prostate carcinogenesis, none of the specific genomic imbalances are mutually exclusive. We therefore propose that the stepwise accumulation of a small subset of genomic imbalances, more than the acquisition of a specific genetic change, causes prostate tumors to become clinically aggressive. The molecular mechanisms behind this nonrandom selection and accumulation of chromosome-level changes remain unknown. Indeed, several candidate genes have already been proposed as relevant targets for the chromosomal regions most affected by gains and losses of genetic material (see Nelson et al., 2003; Kumar-Sinha and Chinnaiyan, 2003, for comprehensive reviews), but many of these lie outside the minimal overlapping regions highlighted in this meta-analysis.

We conclude that clinically aggressive prostate carcinomas tend to display increased genomic complexity resulting from a nonrandom accumulation of a small group of genomic imbalances. More specifically, gain at chromosome arms 7q and 8q and loss at 6q, 10q, and 13q are significantly associated with progression into locally invasive or metastatic disease. Because some of these genomic features seem to predict pathological staging of prostate cancer even when tumor grade is taken into consideration, future studies should evaluate their use as prognostic markers in a preoperative setting using sextant biopsies.

ACKNOWLEDGMENTS

The statistical advice of Dr. Maria José Bento and Dr. Mário Dinis Ribeiro is gratefully acknowledged.

REFERENCES

Affifi A, Azen S. 1979. Statistical analysis: a computer oriented approach. New York: Academic Press. p 318–324.

Aldenderfer MS, Blashfield RK. 1984. Cluster analysis. Sage University paper series on quantitative applications in the social sciences. 07-044. Beverly Hills, CA: Sage.

Alers JC, Krijtenburg PJ, Vis AN, Hoedemaeker RF, Wildhagen MF, Hop WC, van Der Kwast TT, Schroder FH, Tanke HJ, Van Dekken H. 2001. Molecular cytogenetic analysis of prostatic adenocarcinomas from screening studies: early cancers may contain aggressive genetic features. Am J Pathol 158:399–406.

Alers JC, Rochat J, Krijtenburg PJ, Hop WC, Kranse R, Rosenberg C, Tanke HJ, Schroder FH, Van Dekken H. 2000. Identification of genetic markers for prostatic cancer progression. Lab Invest 80:931–942.

Bettendorf O, Schmidt H, Eltze E, Gockel I, Semjonow A, Burger H, Bocker W, Brandt B. 2005. Cytogenetic changes and loss of heterozygosity in atypical adenomatous hyperplasia, in carcinoma of the prostate and in non-neoplastic prostate tissue using comparative genomic hybridization and multiplex-PCR. Int J Oncol 26: 267–274.

Cher ML, Bova GS, Moore DH, Small EJ, Carroll PR, Pin SS, Epstein JI, Isaacs WB, Jensen RH. 1996. Genetic alterations in untreated metastases and androgen-independent prostate cancer detected by comparative genomic hybridization and allelotyping. Cancer Res 56:3091–3102.

Chu LW, Troncoso P, Johnston DA, Liang JC. 2003. Genetic markers useful for distinguishing between organ-confined and locally advanced prostate cancer. Genes Chromosomes Cancer 36: 303–312.

Diep CB, Kleivi K, Ribeiro FR, Teixeira MR, Lingjærde OC, Lothe RA. 2005. The order of genetic events associated with colorectal cancer progression deduced from meta-analysis of copy number changes. Genes Chromosomes Cancer, in press.

Doll JA, Zhu X, Furman J, Kaleem Z, Torres C, Humphrey PA, Donis-Keller H. 1999. Genetic analysis of prostatic atypical adenomatous hyperplasia (adenosis). Am J Pathol 155:967–971.

Dong JT, Chen C, Stultz BG, Isaacs JT, Frierson HF, Jr. 2000. Deletion at 13q21 is associated with aggressive prostate cancers. Cancer Res 60:3880–3883.
Dysvik B, Jonassen I. 2001. J-Express: exploring gene expression data using Java. Bioinformatics 17:369–370.
El Gedaily A, Bubendorf L, Willi N, Fu W, Richter J, Moch H, Mihatsch MJ, Sauter G, Gasser TC. 2001. Discovery of new DNA amplification loci in prostate cancer by comparative genomic hybridization. Prostate 46:184–190.
Etzioni R, Penson DF, Legler JM, di Tommaso D, Boer R, Gann PH, Feuer EJ. 2002. Overdiagnosis due to prostate-specific antigen screening: lessons from U.S. prostate cancer incidence trends. J Natl Cancer Inst 94:981–990.
Foster CS, Bostwick DG, Bonkhoff H, Damber JE, van der KT, Montironi R, Sakr WA. 2000. Cellular and molecular pathology of prostate cancer precursors. Scand J Urol Nephrol Supp 205: 19–43.
Fu W, Bubendorf L, Willi N, Moch H, Mihatsch MJ, Sauter G, Gasser TC. 2000. Genetic changes in clinically organ-confined prostate cancer by comparative genomic hybridization. Urology 56:880–885.
Gleason DF, Mellinger GT. 1974. Prediction of prognosis for prostatic adenocarcinoma by combined histological grading and clinical staging. J Urol 111:58–64.
Hermanek P, Hutter RVP, Sobin LH, Wagner G, Wittekind C. 1997. Prostate. In: Illustrated guide to the TNM/pTNM classification of malignant tumors. Hermanek P, Hutter RVP, Sobin LH, Wagner G, Wittekind C, editors. Heidelberg, Germany: Springer–Verlag. p 278–280.
Höglund M, Frigyesi A, Sall T, Gisselsson D, Mitelman F. 2005. Statistical behavior of complex cancer karyotypes. Genes Chromosomes Cancer 42:327–341.
Höglund M, Gisselsson D, Hansen GB, Mitelman F. 2004a. Wilms tumors develop through two distinct karyotypic pathways. Cancer Genet Cytogenet 150:9–15.
Höglund M, Gisselsson D, Hansen GB, White VA, Sall T, Mitelman F, Horsman D. 2004b. Dissecting karyotypic patterns in malignant melanomas: temporal clustering of losses and gains in melanoma karyotypic evolution. Int J Cancer 108:57–65.
Höglund M, Gisselsson D, Mandahl N, Johansson B, Mertens F, Mitelman F, Sall T. 2001. Multivariate analyses of genomic imbalances in solid tumors reveal distinct and converging pathways of karyotypic evolution. Genes Chromosomes Cancer 31:156–171.
Höglund M, Gisselsson D, Sall T, Mitelman F. 2002. Coping with complexity. Multivariate analysis of tumor karyotypes. Cancer Genet Cytogenet 135:103–109.
ISCN. 1995. An international system for human cytogenetic nomenclature. Mitelman F, editor. Basel, Switzerland: S. Karger.
Jemal A, Tiwari RC, Murray T, Ghafoor A, Samuels A, Ward E, Feuer EJ, Thun MJ. 2004. Cancer statistics, 2004. CA Cancer J Clin 54:8–29.
Jeronimo C, Usadel H, Henrique R, Oliveira J, Lopes C, Nelson WG, Sidransky D. 2001. Quantitation of GSTP1 methylation in non-neoplastic prostatic tissue and organ-confined prostate adenocarcinoma. J Natl Cancer Inst 93:1747–1752.
Kallioniemi A, Kallioniemi OP, Sudar D, Rutovitz D, Gray JW, Waldman F, Pinkel D. 1992. Comparative genomic hybridization for molecular cytogenetic analysis of solid tumors. Science 258: 818–821.
Karayi MK, Markham AF. 2004. Molecular biology of prostate cancer. Prostate Cancer Prostatic Dis 7:6–20.
Kasahara K, Taguchi T, Yamasaki I, Kamada M, Yuri K, Shuin T. 2002. Detection of genetic alterations in advanced prostate cancer by comparative genomic hybridization. Cancer Genet Cytogenet 137:59–63.
Kirchhoff M, Gerdes T, Maahr J, Rose H, Bentz M, Dohner H, Lundsteen C. 1999. Deletions below 10 megabasepairs are detected in comparative genomic hybridization by standard reference intervals. Genes Chromosomes Cancer 25:410–413.
Kirchhoff M, Gerdes T, Rose H, Maahr J, Ottesen AM, Lundsteen C. 1998. Detection of chromosomal gains and losses in comparative genomic hybridization analysis based on standard reference intervals. Cytometry 31:163–173.
Kraggerud SM, Szymanska J, Abeler VM, Kaern J, Eknaes M, Heim S, Teixeira MR, Trope CG, Peltomaki P, Lothe RA. 2000. DNA copy number changes in malignant ovarian germ cell tumors. Cancer Res 60:3025–3030.
Kumar-Sinha C, Chinnaiyan AM. 2003. Molecular markers to identify patients at risk for recurrence after primary treatment for prostate cancer. Urology 62(Suppl 1):19–35.
Matsuda K, Matsuyama H, Hara T, Yoshihiro S, Oga A, Kawauchi S, Furuya T, Izumi H, Naito K, Sasaki K. 2004. DNA sequence copy number aberrations in prostate cancers: a comparison of comparative genomic hybridization data between Japan and European countries. Cancer Genet Cytogenet 152:119–123.
Mattfeldt T, Wolter H, Trijic D, Gottfried HW, Kestler HA. 2002. Chromosomal regions in prostatic carcinomas studied by comparative genomic hybridization, hierarchical cluster analysis and self-organizing feature maps. Anal Cell Pathol 24:167–179.
Mitelman F, Johansson B, Mertens F. 2005. Mitelman database of chromosome aberrations in cancer. Available at: http://cgap.nci.nih.gov/Chromosomes/Mitelman.
Nelson WG, De Marzo AM, Isaacs WB. 2003. Prostate cancer. N Engl J Med 349:366–381.
Nupponen NN, Kakkola L, Koivisto P, Visakorpi T. 1998. Genetic alterations in hormone-refractory recurrent prostate carcinomas. Am J Pathol 153:141–148.
Paris PL, Albertson DG, Alers JC, Andaya A, Carroll P, Fridlyand J, Jain AN, Kamkar S, Kowbel D, Krijtenburg PJ, Pinkel D, Schroder FH, Vissers KJ, Watson VJ, Wildhagen MF, Collins C, Van Dekken H. 2003. High-resolution analysis of paraffin-embedded and formalin-fixed prostate tumors using comparative genomic hybridization to genomic microarrays. Am J Pathol 162: 763–770.
Rokman A, Koivisto PA, Matikainen MP, Kuukasjarvi T, Poutiainen M, Helin HJ, Karhu R, Kallioniemi OP, Schleutker J. 2001. Genetic changes in familial prostate cancer by comparative genomic hybridization. Prostate 46:233–239.
Schulz WA, Elo JP, Florl AR, Pennanen S, Santourlidis S, Engers R, Buchardt M, Seifert HH, Visakorpi T. 2002. Genomewide DNA hypomethylation is associated with alterations on chromosome 8 in prostate carcinoma. Genes Chromosomes Cancer 35:58–65.
Stanford JL, Feng Z, Hamilton AS, Gilliland FD, Stephenson RA, Eley JW, Albertsen PC, Harlan LC, Potosky AL. 2000. Urinary and sexual function after radical prostatectomy for clinically localized prostate cancer: the Prostate Cancer Outcomes Study. JAMA 283:354–360.
Steineck G, Helgesen F, Adolfsson J, Dickman PW, Johansson JE, Norlen BJ, Holmberg L. 2002. Quality of life after radical prostatectomy or watchful waiting. N Engl J Med 347:790–796.
Steiner T, Junker K, Burkhardt F, Braunsdorf A, Janitzky V, Schubert J. 2002. Gain in chromosome 8q correlates with early progression in hormonal treated prostate cancer. Eur Urol 41:167–171.
Teixeira MR, Ribeiro FR, Eknaes M, Waehre H, Stenwig AE, Gierecksky KE, Heim S, Lothe RA. 2004. Genomic analysis of prostate carcinoma specimens obtained via ultrasound-guided needle biopsy may be of use in preoperative decision-making. Cancer 101:1786–1793.
Teixeira MR, Waehre H, Lothe RA, Stenwig AE, Pandis N, Gierecksky KE, Heim S. 2000. High frequency of clonal chromosome abnormalities in prostatic neoplasms sampled by prostatectomy or ultrasound-guided needle biopsy. Genes Chromosomes Cancer 28:211–219.
Van Dekken H, Alers JC, Damen IA, Vissers KJ, Krijtenburg PJ, Hoedemaeker RF, Wildhagen MF, Hop WC, van der Kwast TH, Tanke HJ, Schroder FH. 2003. Genetic evaluation of localized prostate cancer in a cohort of forty patients: gain of distal 8q discriminates between progressors and nonprogressors. Lab Invest 83:789–796.
Visakorpi T, Kallioniemi AH, Syvänen AC, Hyytinen ER, Karhu R, Tammela T, Isola JJ, Kallioniemi OP. 1995. Genetic changes in primary and recurrent prostate cancer by comparative genomic hybridization. Cancer Res 55:342–347.
Wolter H, Gottfried HW, Mattfeldt T. 2002a. Genetic changes in stage pT2N0 prostate cancer studied by comparative genomic hybridization. BJU Int 89:310–316.
Wolter H, Trijic D, Gottfried HW, Mattfeldt T. 2002b. Chromosomal changes in incidental prostatic carcinomas detected by comparative genomic hybridization. Eur Urol 41:328–334.