

# Analysis of molecular intra-patient variation and delineation of a prognostic 12-gene signature in non-muscle invasive bladder cancer; technology transfer from microarrays to PCR

L Dyrskjöt<sup>\*1</sup>, T Reinert<sup>1</sup>, A Novoradovsky<sup>2</sup>, TCM Zuiverloon<sup>3</sup>, W Beukers<sup>4</sup>, E Zwarthoff<sup>4</sup>, N Malats<sup>5</sup>, FX Real<sup>5</sup>, U Segersten<sup>6</sup>, P-U Malmström<sup>6</sup>, M Knowles<sup>7</sup>, C Hurst<sup>7</sup>, J Sorge<sup>2</sup>, M Borre<sup>8</sup> and TF Ørntoft<sup>1</sup>

<sup>1</sup>Department of Molecular Medicine, Aarhus University Hospital, Brendstrupgaardsvej 100, Skejby, Aarhus N, 8200 Aarhus, Denmark; <sup>2</sup>Decisive Diagnostics, Jackson, WY 83002, USA; <sup>3</sup>Department of Urology, Erasmus Medical Center, Rotterdam 3000 DR, The Netherlands; <sup>4</sup>Department of Pathology, Erasmus Medical Center, Rotterdam 3000 DR, The Netherlands; <sup>5</sup>Centro Nacional de Investigaciones Oncológicas (CNIO), Madrid 28029, Spain; <sup>6</sup>Department of Surgical Sciences, Urology, Academic Hospital, Uppsala 75185, Sweden; <sup>7</sup>Cancer Research UK Clinical Centre, St James's University Hospital, Leeds LS9 7TF, UK; <sup>8</sup>Department of Urology, Aarhus University Hospital, Brendstrupgaardsvej 100, Skejby, Aarhus N, 8200 Aarhus, Denmark

**BACKGROUND:** Multiple clinical risk factors and genetic profiles have been demonstrated to predict progression of non-muscle invasive bladder cancer; however, no easily clinical applicable gene signature has been developed to predict disease progression independent of disease stage and grade.

**METHODS:** We measured the intra-patient variation of an 88-gene progression signature using 39 metachronous tumours from 17 patients. For delineation of the optimal quantitative reverse transcriptase PCR panel of markers, we used 115 tumour samples from patients in Denmark, Sweden, UK and Spain.

**RESULTS:** Analysis of intra-patient variation of the molecular markers showed 71% similar classification results. A final panel of 12 genes was selected, showing significant correlation with outcome. In multivariate Cox regression analysis, we found that the 12-gene signature was an independent prognostic factor (hazard ratio = 7.4 (95% confidence interval: 3.4–15.9),  $P < 0.001$ ) when adjusting for stage, grade and treatment. Independent validation of the 12-gene panel and the determined cut-off values is needed and ongoing.

**CONCLUSION:** Intra-patient marker variation in metachronous tumours is present. Therefore, to increase test sensitivity, it may be necessary to test several metachronous tumours from a patient's disease course. A PCR-based 12-gene signature significantly predicts disease progression in patients with non-muscle invasive bladder cancer.

*British Journal of Cancer* (2012) **107**, 1392–1398. doi:10.1038/bjc.2012.412 www.bjcancer.com

Published online 13 September 2012

© 2012 Cancer Research UK

**Keywords:** bladder cancer; PCR; heterogeneity; outcome; progression

A total of 70 530 new bladder cancer cases and 14 680 deaths were estimated in the United States alone in 2010 (Jemal *et al*, 2010). About 75% of patients present with non-muscle invasive tumours were mostly treated with a local, organ-sparing approach. The remaining 25% of patients are initially diagnosed with muscle invasive cancers, most often requiring immediate and more radical treatment. More than 60% of patients with non-muscle invasive bladder cancer experience tumour recurrences and around 20% of the patients develop disease progression to a muscle invasive bladder cancer (Millan-Rodriguez *et al*, 2000; Sylvester *et al*, 2006). Clinical risk factors associated with a high risk of disease progression to a muscle invasive cancer include deep invasion of the lamina propria, high-grade tumour, large tumour size, concurrent carcinoma *in situ* (CIS), tumour multiplicity and recurrence of high-risk non-muscle invasive tumours (Hermann *et al*, 1998). The clinical risk factors cannot predict the individual disease course accurately, and currently no molecular markers are

available for clinical use for predicting tumour recurrence or later disease progression (Ehdaie and Theodorescu, 2008).

Several microarray-based gene-expression signatures for diagnosis or prognosis have been identified in many cancers, including bladder cancer (Rosenwald *et al*, 2002; van de Vijver *et al*, 2002; van't Veer *et al*, 2002; Ramaswamy *et al*, 2003; Dyrskjot *et al*, 2007a). Although clinically applicable, it may, however, be easier to implement tests for clinical use when simpler, cheaper and more sensitive methods are being used. We have previously identified a molecular gene-expression signature for predicting disease progression in patients with non-muscle invasive bladder cancer (Dyrskjot *et al*, 2005). The progression signature was validated successfully in a retrospective study that included 294 patients from five different countries. The molecular signature was significantly correlated with progression-free survival ( $P < 0.001$ ) and it was shown to be significantly associated with disease progression after adjustment for age, sex, stage, grade and treatment (hazard ratio (HR) = 2.3;  $P = 0.007$ ) (Dyrskjot *et al*, 2007b). One of the objectives to be addressed before clinical implementation is intra-patient variation of the molecular signature when applied to multiple tumours from the same patient. How often do we need to test in each patient? Another objective is the generation of a clinically applicable gene

\*Correspondence: Dr L Dyrskjöt; E-mail: lars@ki.au.dk

Received 7 June 2012; revised 10 August 2012; accepted 14 August 2012; published online 13 September 2012

signature, and this requires transferring the markers from the microarray platform to a robust PCR-based platform.

Here we address the intra-patient variation of the original 88-gene signature by analysing 39 metachronous tumours from 17 patients. We combined this analysis with mutation screening of the genes frequently mutated in bladder cancer: *FGFR3* (16–74% mutated – dependent on stage), *PIK3CA* (13–27% mutated) and *RAS* family (*NRAS*, *KRAS* and *HRAS*; 13% mutated) (Billerey *et al*, 2001; Kompier *et al*, 2010). Then, we describe the transfer of the optimal signature genes from the microarray platform, plus additional prognostic genes identified, to a PCR platform for easier clinical implementation of the molecular classifier.

## MATERIALS AND METHODS

### Patients and tumour samples

The majority of the patients and tumour samples used in this work were used previously for validating our microarray-based 88-gene progression signature (Dyrskjot *et al*, 2007b). Urothelial carcinomas were taken from patients that were operated in the years 1987–2000 in hospitals in Denmark, Sweden, Spain and England and frozen immediately at  $-80^{\circ}\text{C}$ . All patients were followed according to the valid guidelines at patient inclusion time. Informed written consent was obtained from all patients, and research protocols were approved by the institutional review boards or ethical committees in all involved countries. Diagnostic pathology slides were reevaluated according to the World Health Organisation (WHO) 2004 guidelines. However, for 19 out of the 115 patients, it was not possible to acquire the diagnostic sections for review and consequently the original grading was translated into the WHO 2004 grading system (G1 + G2 = low grade, G3 + G4 = high grade), although this translation is too simplified in some cases. Progression-free survival time was recorded from sampling visit and censored at the time of the last control cystoscopy or at cystectomy. Progression of the disease was defined as (i) histologically confirmed invasion into the bladder muscle or (ii) cases where the patient died of bladder cancer without a verified progression event as described above ( $n=5$ ). In case of bladder cancer death, survival was recorded from the date at which tissue was collected and until the last annotation of the patient being alive.

### RNA extraction and cDNA generation

Total RNA was extracted from the Danish, and English samples using a standard Trizol RNA extraction method (Invitrogen, Carlsbad, CA, USA). Total RNA from Swedish and Spanish samples was extracted using RNeasy mini kit (Qiagen, Valencia, CA, USA). All RNA was quality controlled using an Agilent Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA) (criteria: 28S/18S > 1 and RIN > 5). In all, 1  $\mu\text{g}$  of total RNA was DNase-treated using DNase I Amplification Grade (Invitrogen) to avoid amplification of genomic DNA. Adequate DNase treatment was verified in all samples by quantitative reverse transcriptase PCR (qRT-PCR) using intron spanning *GAPDH* primers. The DNase-treated total RNA was converted to cDNA using oligo (dT) priming and SuperScript II Reverse Transcriptase (Invitrogen) according to the standard protocols.

### Initial selection of 96 genes for analysis

Genes were selected using two different approaches. First, genes were selected from the prognostic gene signatures published previously (Dyrskjot *et al*, 2007b). Second, we reanalysed previously published microarray data (Dyrskjot *et al*, 2003, 2004; Aaboe *et al*, 2005; Dyrskjot *et al*, 2005). For each gene, the microarray hybridisation intensity was plotted against progression-free survival of each patient and linear and quadratic regression equations were computed. The noise was calculated as a function of squared deviations of each patient from the squared regression line. Then, the ratio of linear regression coefficient to noise was calculated for

each gene and used as a measure of ‘signal-to-noise’ ratio. Genes were then selected based on ‘signal-to-noise’ performance. The two different approaches showed large overlap in genes selected, and we selected those showing the best performance with each approach.

### Primer design and PCR assays

Microarray layout files were used to map the Affymetrix (Santa Clara, CA, USA) probe sets to mRNA sequences and this region was then used to design primers for qRT-PCR assays. For each of the selected candidate genes, we designed three primer pairs using PRIMER3 software (Rozen and Skaletsky, 2000). For sensitivity measures, we used 1, 0.1 and 0.01 ng cDNA template. A minimum of two replicates had to amplify at a particular concentration to qualify as the limit of sensitivity. Furthermore, the threshold cycle (Ct) spread between the replicates at this concentration had to be less than one cycle. Specificity measures were performed using no template control measurements where Ct should be > 40. The qRT-PCR using the final primer set was carried out using a 7900HT Fast Real-Time PCR System (Life Technologies, Carlsbad, CA, USA) in 384-well plates. A total of 115 tumour samples were analysed. All reactions were performed in triplicate in 10- $\mu\text{l}$  volumes using SYBR Green PCR Master Mix (Life Technologies) and the standard protocols were applied.

### Normalisation, optimal 12-gene set selection and molecular classification

To identify the best-performing candidate genes, we normalised data using average total-patient Ct values and using an average global total. Classifier genes were selected based on different statistical methods applied to normalised Ct values and clinical outcome (progression vs no progression): Pearson’s correlation,  $P<0.01$ ; Wilcoxon signed-rank test,  $P<0.01$ ; Cox regression analysis,  $P<0.01$ ; and area under curve (AUC) from receiver operating characteristics (ROCs),  $\text{AUC}>0.65$ . Classifier genes with high correlation were not excluded from the final signature. Computation of risk scores utilised non-normalised Ct values based on the formula: average (Ct (genes downregulated in progressing tumours) – average (Ct (genes upregulated in progressing tumours))). This approach eliminated the requirement of using normalised Ct values. The optimal cut-off value (risk score with the highest accuracy) and the 90% sensitivity cut-off value were identified from ROC curves, and risk scores were then dichotomised based on these two different cut-off values. Dichotomised signature values were used to classify samples as either high or low risk of progression (1 = high risk and 0 = low risk).

### Affymetrix Exon ST 1.0 array analysis and classifier construction

Affymetrix human Exon ST 1.0 arrays were used for measuring gene expression in 39 tumours from 17 patients as described previously (Thorsen *et al*, 2008). *itPLIER* normalisation and generation of gene-expression measures were performed in GeneSpring GX 10.0 with no baseline adjustment. To identify the original 88-gene signature, we used BLASTN (NCBI) against human RefSeq RNA sequences for all original probe sequences to updated Gene Symbol annotation. Following, we used the Gene Symbols for identifying the corresponding probe sets on the Exon ST 1.0 arrays. We preferentially selected ‘core’ probe set with the highest number of exon clusters. A maximum likelihood classifier was constructed as described previously (Dyrskjot *et al*, 2003) using 10 tumours that were originally used for building the classifier and for testing the 39 independent tumour samples.

### Mutation analysis

Mutation analysis of the genes *FGFR3*, *PIK3CA*, *NRAS*, *KRAS* and *HRAS* was performed essentially as described previously (Kompier

**Table 1** Intra-patient variation based on 88-gene classifier results and mutation analysis

Sample	Stage	Grade <sup>a</sup>	88-gene classifier <sup>b</sup>	Mutation analysis				
				FGFR3	PIK3CA	NRAS	HRAS	KRAS
532-12	Ta	II	I	S249C	WT	WT	WT	WT
532-15	Ta	II	I	S249C	WT	WT	WT	WT
1888-1	Ta	II	I	S249C	WT	WT	WT	WT
1888-2	Ta	II	I	S249C	WT	WT	WT	WT
1841-4	Ta	II	I	S249C	WT	WT	WT	WT
1841-6	Ta	II	I	S249C	WT	WT	WT	WT
1926-1	Ta	II	I	G372C	WT	WT	WT	WT
1926-4	Ta	II	I	G372C	WT	WT	WT	WT
1926-6	Ta	II	I	G372C	WT	WT	WT	WT
1657-1	Ta	I	I	S249C	WT	WT	WT	WT
1657-3	Ta	II	I	S249C	WT	WT	WT	WT
1899-2	Ta	III	<b>I</b>	<b>G372C</b>	WT	WT	WT	WT
1899-3	T2-4	III	<b>2<sup>c</sup></b>	<b>WT</b>	WT	WT	WT	WT
1992-1	Ta	I	<b>2</b>	S249C	WT	WT	WT	WT
1992-3	T1	II	<b>I</b>	S249C	WT	WT	WT	WT
2010-1	T1	III	2	S249C	WT	WT	WT	WT
2010-4	T2-4	III	2	S249C	WT	WT	WT	WT
1375-1	T1	II	I	S249C	E545K	WT	WT	WT
1375-7	Ta	II	I	S249C	E545K	WT	WT	WT
1735-1	T1	III	2	WT	WT	WT	WT	WT
1735-4	Ta	II	2	WT	WT	WT	WT	WT
1702-1	Ta	I	I	S249C	WT	WT	WT	WT
1702-7	Ta	II	I	S249C	WT	WT	WT	WT
1702-11	Ta	I	I	S249C	WT	WT	WT	WT
1790-1	Ta	I	I	Y375C	WT	WT	WT	WT
1790-3	Ta	II	I	Y375C	WT	WT	WT	WT
562-3	Ta	I	2	Y375C	WT	WT	WT	WT
562-5	Ta	II	2	Y375C	WT	WT	WT	WT
1521-1	Ta	I	I	S249C	WT	WT	WT	WT
1521-6	Ta	I	I	S249C	WT	WT	WT	WT
1521-10	T1	II	I	—	—	—	—	—
1732-1	T1	II	<b>2</b>	—	—	—	—	—
1732-11	Ta	II	<b>I</b>	S249C	WT	WT	WT	WT
1732-12	T2-4	III	<b>2</b>	—	—	—	—	—
1829-1	Ta	II	<b>I</b>	S249C	WT	WT	WT	WT
1829-2	Ta	II	<b>2</b>	S249C	WT	WT	WT	WT
1829-3	T1	III	<b>I</b>	S249C	WT	WT	WT	WT
2363-1	T1	III	<b>I</b>	<b>S249C</b>	WT	WT	WT	WT
2363-2	T2-4	III	<b>2</b>	<b>WT</b>	WT	WT	WT	WT

<sup>a</sup>Bergkvist grading. <sup>b</sup>I = classified as a non-progressing tumour and 2 = classified as a progressing tumour. <sup>c</sup>Less than 5% difference in classification distance between the two classes. Bold denotes intra-patient differences in 88-gene classifier results or in mutation status.

et al, 2010) for 36 out of the 39 tumours samples analysed on the Exon ST 1.0 microarray platform.

### Statistical procedures

We used STATA 10.0 (Stata Corporation, College Station, TX, USA) for statistical analysis. Variables with a *P*-value < 0.05 in univariate analysis were included in multivariate analysis. The assumptions of proportional hazards were verified.

## RESULTS

### Intra-patient variation

We set out to analyse the progression signature in 39 metachronous bladder tumours from 17 patients by Exon ST 1.0 microarray analysis to measure the intra-patient variation of the signature. The mutational status of the samples was also analysed to obtain an indication of the genetic relatedness of the metachronous tumours. The classification result and mutation analysis of the genes *FGFR3*, *PIK3CA*, *NRAS*, *KRAS* and *HRAS* are shown in Table 1. *NRAS*, *HRAS* and *KRAS* were not mutated in any of the tumours analysed. Multiple tumours from 12 patients (71%) showed similar intra-patient classification results, however, in

another five cases the intra-patient analysis results were dissimilar; two of these patients showed variation in mutation status of the *FGFR3* gene, indicating differences in tumour clonality. All other tumours showed similar mutation status.

### Patient and marker selection

Initially, we selected 96 genes associated with disease aggressiveness from published gene signatures (38 genes) and from reanalysis of microarray data (58 genes). Following primer design and initial validation of primer sensitivity and specificity, we decreased the number of interesting genes to 35 in total (Table 2). We performed qRT-PCR measurements of the 35 genes using 115 tumour samples from patients in Denmark, Sweden, UK and Spain. Previously, 102 of these tumours were used for validating our microarray-based 88-gene signature. Patients with no disease progression were followed for a minimum of 60 months (the median follow-up time was 88 months). Clinical and histopathological characteristics for all patients are shown in Table 3.

### 12-gene PCR signature

We selected the genes for inclusion in the optimal PCR signature by comparing normalised Ct values and clinical outcome using

**Table 2** Primer sequences for the 35 candidate genes

Gene target <sup>a</sup>	Upstream primer sequence 5'–3'	Downstream primer sequence 5'–3'	Gene selection <sup>b</sup>
ADAM10	5'-CAGTACTTATGGAATTGCTCTG-3'	5'-TTGGATTACTTGGAGTATGAACAC-3'	1
AURKB	5'-TTTGCTATGAGCTGCTGGTG-3'	5'-ACTTTAGTCCACCTTGACGATG-3'	1
<b>BIRC5</b>	5'-CTGAAGTCTGGCGTAAGATGATG-3'	5'-GAAGCTGTAACAATCCACCCTG-3'	1
C10orf58	5'-GTAACCTACTTTCTGTTCTGGAAAGC-3'	5'-TTTTCTCTGAGGCCAAAGTCTG-3'	1
CDC20	5'-AGTCCAATGTCTGGCAACAG-3'	5'-CCAGAGCACACATTCCAGATG-3'	1
<b>CDC25B</b>	5'-GATGGAAGTTGGATGGATG-3'	5'-ACCTGGTTGGGTATGCAAG-3'	1
<b>COL4A1</b>	5'-CTGCCTGGAGGATTTAGAAGTG-3'	5'-CTGTAAGCGTTTGCCTAGTAATTG-3'	1
<b>FABP4</b>	5'-AGAGAAAACGAGAGGATGATAAACTG-3'	5'-CTTATGCTCTCATAAACTCTCGT-3'	1
IGF2	5'-CATCGTTGAGGAGTGCTGTTTC-3'	5'-GGGTAGCACAGTACGTCTCCAG-3'	1
ITGB4	5'-CATCATCCCTGACATCCCTATC-3'	5'-GTAGAACGTCATCGCTGTACATAAG-3'	1
<b>KPNA2</b>	5'-GCAGATTTTAAAGACACAAAAGGAAG-3'	5'-AAGGTACACAATCTGTTCAACTGTTTC-3'	1
MAT2B	5'-TTGCTAAAGAACTAAAGGGCAGTC-3'	5'-AGTTTAGCCAGGACAAAATAAG-3'	1
<b>MBNL2</b>	5'-ACTTCATCCAGTGCCTCACTTTC-3'	5'-GGGGTTACAGGTGCTAGGTAAGG-3'	1
MCM7	5'-GAGATGTCAAAGGACTCTTTCTAGG-3'	5'-GCAATATCACATCTGCTGGTC-3'	1
<b>MSN</b>	5'-CCTGACCTTGAGGACTCTTGTG-3'	5'-AATATAGGCATATCACCAAGTGAGC-3'	1
PPP2R5C	5'-GTACTACATTGAAAATAAACCGGTGAC-3'	5'-TACATTTGGAAAGAGTGAAGATGC-3'	1
TCF4	5'-GAATCACATGGGACAGATGAAAAG-3'	5'-AATACAGCTGTTAAGGAAGTGGTCTC-3'	1
ACTA2	5'-TCTCTAGCACACAACCTGGAATGTC-3'	5'-CTAGGAATGATTTGGAAAAGCACTG-3'	2
CDH5	5'-AAACAATTCCTGTAACCTTCTATTTTC-3'	5'-CTTGTCATGCACCAGTTTGG-3'	2
CDKN3	5'-ATCTCTACCAGCAATGTGGAATTATC-3'	5'-CTATGTCAGGAGTCCCTCCATC-3'	2
<b>COL18A1</b>	5'-GGGCTGGTCTGTAATTGTGTG-3'	5'-AAAAGGTCACCTTTATTTGCTGTC-3'	2
<b>COL4A3BP</b>	5'-TTTCTGTGGATCATGACAGTGC-3'	5'-CAAGGTTTGACAAATCATAGCAAC-3'	2
CPS1	5'-GGAAGTAAGGTTCACTCCCTTAAGAC-3'	5'-CCTTACAGTGGGTGGAATTAATAGTG-3'	2
DCTD	5'-TGTGCCCTTCTCTTAATCTC-3'	5'-GAAAGCCTTTTCTCAACACAGG-3'	2
IER2	5'-CTTGCCAGGGAGTTTCTGAG-3'	5'-ATTTCTAACAAAACGCCAGGTAGAC-3'	2
LBR	5'-ATCAGAAAGTGGTGGCGTTTTC-3'	5'-TTACCAGGGAAGAATTTAATGTCC-3'	2
LGALS1	5'-CTGAATCTCAAACCTGGAGAGTG-3'	5'-GGTTCAGCAGGAGCTCTTAGC-3'	2
<b>NEK1</b>	5'-CTAAAAGACCAGTCTCAGGACAAAAC-3'	5'-CTAAAGGATTCCATTTAGCGGC-3'	2
NR1H3	5'-GGAATTCATCAACCCCATCTTC-3'	5'-GATAGCAATGAGCAAGGCAAC-3'	2
PEA15	5'-ACTCCTTATATTGCTGTGAGATTGC-3'	5'-ACCTTTATTCCGGGTTAGAACAAG-3'	2
<b>SKAP2</b>	5'-TGGAGATGATGATTTTGGAGTCC-3'	5'-CTAAATCCAAAGCATTTGCAGAC-3'	2
SDC1	5'-AGACACCTGGACATCCTCCTC-3'	5'-TAAGCAAGTAAGTGCAGGAGCC-3'	2
SECI4L1	5'-TGTTTCTACCTTTAGTACCTTGCCAC-3'	5'-AGTACTAAGAAATGGGAAATGACAGC-3'	2
<b>UBE2C</b>	5'-TCTAGGAGAACCAACATTGATAGTC-3'	5'-TCTTGCAGGTACTTCTAAAAGCTG-3'	2
WNT2B	5'-ATAAGAACTGTGCAAGCTCCC-3'	5'-TCTACTCTCCCTCAAATCTCCAG-3'	2

<sup>a</sup>Genes in bold are included in the final 12-gene signature. <sup>b</sup>Gene selection method: 1 = from previously published gene signatures and 2 = from reanalysis of gene-expression data.

**Table 3** Clinical and histopathological characteristics for all 115 patients

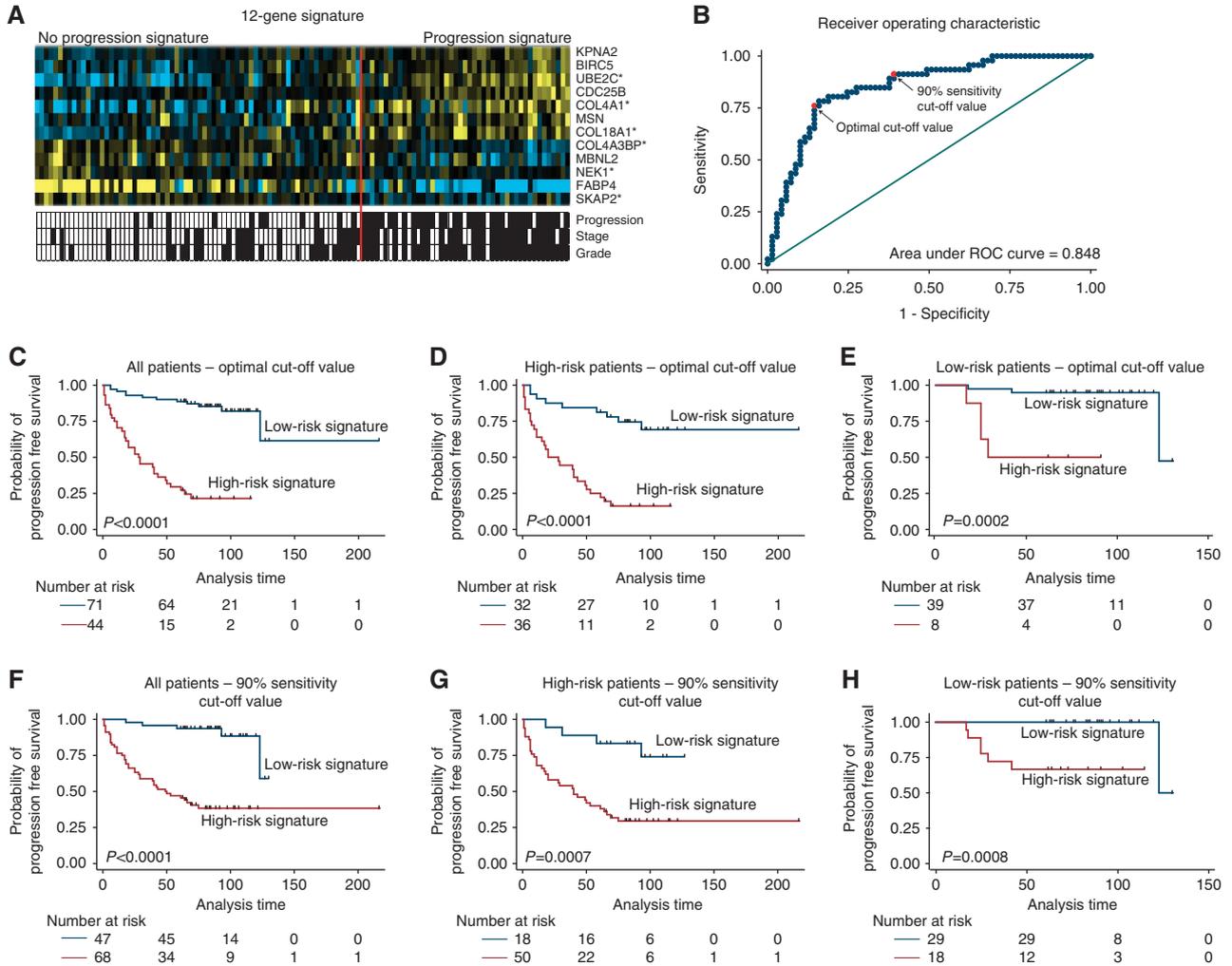
Number of patients	115
Median follow-up time in months for all patients (range)	69 (1–216)
Median follow-up time in months for progressing patients (range)	20 (1–123)
Median follow-up time in months for non-progressing patients (range)	88 (61–216)
Median age (range)	70 (33–88)
Male–female ratio	4.9
Stage	
Ta	62 (54%)
T1	53 (46%)
Grading (WHO 2004)	
PUNLMP	17 (15%)
Low grade	33 (29%)
High grade	65 (56%)
Concomitant CIS	
Yes	19 (17%)
No	50 (43%)
Unspecified	46 (40%)
Adjuvant therapy (BCG or MMC)	
Yes	34 (30%)
No	81 (70%)
Number of progression events to stage T2–4 bladder cancer	
Ta	13 (30%)
T1	33 (70%)

Abbreviations: BCG = bacillus Calmette-Guérin; CIS = concurrent carcinoma *in situ*; MMC = mitomycin C; WHO = World Health Organisation.

Pearson's correlations, ROC analysis, Wilcoxon signed-rank tests and Cox regression analysis. Using these different statistical methods in a combined approach, we selected 12 genes that showed significant correlation with outcome (see Figure 1A and Table 4); seven were upregulated in progressing tumours compared with non-progressing tumours (**KPNA2**, **BIRC5**, **UBE2C**\*, **CDC25B**, **COL4A1**\*, **MSN** and **COL18A1**\*) and five were downregulated in progressing tumours compared with non-progressing tumours (**COL4A3BP**\*, **MBNL2**, **NEK1**\*, **FABP4** and **SKAP2**\*). The six genes marked with an asterisk were not previously included in our gene signature for outcome prediction. Selecting fewer genes using more stringent selection criteria did not result in increased classification performance (results not shown). A schematic representation of the steps involved in reaching the final 12-gene PCR signature is shown in Figure 2.

### Molecular classification

Molecular prediction of outcome was carried out for all 115 patients based on the generated PCR risk scores. An optimal cut-off value (0.79; 76% sensitivity and 86% specificity) was identified using ROC analysis (Figure 1B) and the dichotomised scores showed significant correlation with progression-free survival ( $P < 0.0001$ , log-rank test; Figure 1C). In multivariate Cox regression analysis, we found that the 12-gene signature was an independent prognostic variable (HR = 7.4 (95% confidence interval: 3.4–15.9),  $P < 0.001$ ) when adjusting for stage, grade and BCG/MMC treatment (Table 5). When only including high-risk patients in the analysis (i.e., stage T1, high grade or concomitant



**Figure 1** The 12-gene PCR signature for predicting disease progression. **(A)** Expression patterns of the 12 genes in the 115 tumour samples analysed. The upper seven genes are upregulated in progressing tumours and the lower five genes are downregulated in progressing tumours. Yellow, upregulation of the gene; blue, downregulation; and black, median expression. The vertical red line separates tumours classified with the progression signature (right) from tumours with the non-progression signature (left) based on the optimal cut-off value. The black boxes below the heat map denote progressing tumours, stage T1 tumours and high-grade tumours. **(B)** ROC curve for the 12-gene signature values (average Ct (genes downregulated in progressing tumours) – average Ct (genes upregulated in progressing tumours)). The value with the highest accuracy is indicated as the optimal cut-off value for the dichotomised test (used in **C**, **D** and **E**), and the 90% sensitivity value is indicated (used in **F**, **G** and **H**). **(C, F)** Kaplan–Meier plots of progression-free survival as a function of the 12-gene PCR signature. All 115 patients are included in the analysis. **(D, G)** Kaplan–Meier plots of progression-free survival as a function of the 12-gene PCR signature when including high-risk patients only (stage T1 or high grade or CIS). **(E, H)** Kaplan–Meier plots of progression-free survival as a function of the 12-gene PCR signature when including low-risk patients only.

CIS), we also found a significant correlation with progression-free survival ( $P < 0.0001$ , log-rank test; Figure 1D). Similarly, when including only the low-risk subpopulation in the analysis, we also found a significant correlation with progression-free survival ( $P = 0.0002$ , log-rank test; Figure 1E). We calculated the predictive accuracy using Harrell's C to 73% for clinical variables only, 75% for the 12-gene signature, and this number was raised to 82% when including clinical variables and the 12-gene signature in the model. Consequently, accurate prediction of progression is increased when including the molecular markers. When changing the cut-off value to  $-0.17$  to ensure a 90% sensitivity (and 61% specificity) of the test, we also observed significant correlations with outcome for the entire patient cohort and the risk subgroups as shown in Figure 1F–H.

## DISCUSSION

Identification of highly sensitive and specific biomarkers for prediction of outcome for patients with non-muscle invasive

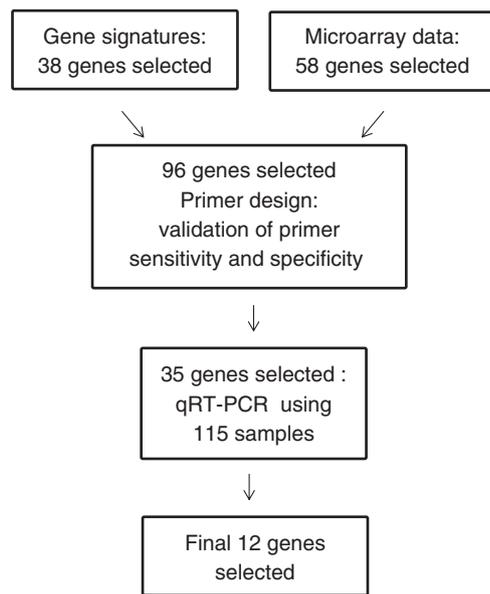
bladder tumours is of utmost clinical importance. Here we report the required step in clinical implementation of gene-expression signatures; the generation of a 12-gene signature for disease progression using qRT-PCR. We showed that the final 12-gene signature correlated highly significantly with outcome, and showed independent prognostic value when stratifying for stage, grade and treatment. Furthermore, by mutational analysis of genes frequently mutated in bladder cancer (*FGFR3*, *PIK3CA* and *RAS* family genes) combined with microarray-based analysis of an earlier reported 88-gene prognostic signature, we showed that intra-patient reproducibility and switch in tumour clonality may indicate that the 12-gene PCR signature analysis of several recurrent tumours is necessary in each single patient.

The analysis of intra-patient reproducibility of the microarray-based 88-gene molecular signature showed that in most cases (71%), identical classifications were obtained. In two of the five cases where divergent results were obtained, we also observed differences in the mutation status of *FGFR3*, indicating that the tumours may have arisen from different subclones.

**Table 4** Progression-associated genes selected for the final 12-gene PCR signature and associated statistics based on 115 tumour samples

GENE	Pearson's correlation <sup>a</sup>	Wilcoxon signed-rank test P-value	Cox regression analysis		ROC AUC
			Beta coefficient	P-value	
<i>CDC25B</i>	-0.268	0.004	-0.605	0.002	0.665
<i>KPNA2</i>	-0.234	0.004	-0.490	0.006	0.663
<i>BIRC5</i>	-0.251	0.004	-0.449	0.005	0.667
<i>COL18A1</i>	-0.307	<0.001	-0.368	0.002	0.704
<i>MSN</i>	-0.277	<0.001	-0.328	0.007	0.690
<i>UBE2C</i>	-0.331	<0.001	-0.315	<0.001	0.729
<i>COL4A1</i>	-0.345	<0.001	-0.300	<0.001	0.740
<i>FABP4</i>	0.374	<0.001	0.193	<0.001	0.729
<i>MBNL2</i>	0.297	<0.001	0.459	0.001	0.707
<i>SKAP2</i>	0.270	0.004	0.566	0.004	0.664
<i>COL4A3BP</i>	0.359	<0.001	0.777	<0.001	0.746
<i>NEK1</i>	0.253	0.003	0.804	0.003	0.671

Abbreviations: AUC = area under curve; ROC = receiver operating characteristic.  
<sup>a</sup>P < 0.01 if the Pearson's correlation is > 0.176 or < -0.176. A negative value denotes genes upregulated in progressing tumours.

**Figure 2** Schematic overview of steps involved in delineation of the final 12-gene PCR signature.

The monoclonal and oligoclonal nature of bladder tumours has been observed in several previous reports (Hartmann *et al*, 2000; Zieger *et al*, 2005). Sometimes the latest removed tumour was classified as high risk – sometimes it was the first removed tumour. This seems controversial; however, previous publications have indicated that the order of tumour resection may not be the order of formation of the tumour in the bladder. Very early tumours that grow very slowly may be resected later than offspring from this tumour that has a more rapid proliferation (van Tilborg *et al*, 2000). Whether a switch in clonality is reflected at cystoscopic examination is at present unknown, but could be triggering a new molecular analysis. The analysis of intra-patient, and probably intra-tumour, variation of molecular signatures may be highly relevant in other cancer diseases with synchronous multifocal presentation like, for example, breast and prostate cancer.

Importantly, most of the samples used in this study to delineate the 12-gene PCR signature were analysed in the previous microarray-based validation study also, and consequently this study is to a large degree a technology-transfer study, and not a true validation of the molecular signature. Therefore, the relatively high HRs reported here for the 12-gene signature are based on the optimal cut-off values, and hence represent best case values, and the proposed 12-gene signature needs external validation before eventual clinical implementation. Furthermore, the technical reproducibility of the 12-gene signature needs to be further studied; however, because there is no need for normalisation genes, the assay and associated statistical calculations are easily handled.

The population studied here is neither consecutive nor included within a certain time interval, and consequently different biases may have been introduced; different follow-up guidelines and treatment regimens have been applied in different countries and in different time periods (e.g., a large fraction of high-risk patients have not received adjuvant treatment, and diagnosis of CIS from selected site biopsies has not been performed routinely). Furthermore, few women were included and several parameters regarding, for example, smoking history were not available. Therefore, prospective studies are needed for further validation of the findings. To address this, we are currently validating the 12-gene signature in an ongoing prospective study that will ultimately include 1200 patients with bladder cancer (UROMOL, FP7 EU project).

The 12-gene signature selected in this work includes six novel prognostic markers not included in previously reported signatures from our group. *UBE2C* and *COL18A1* have been identified earlier in another meta-analysis of bladder cancer gene-expression data (Oncomine). The *UBE2C* and *COL18A1* expression was described as being upregulated in progressing tumours and both genes were included in a 57-gene qRT-PCR signature (Wang *et al*, 2009). *SKAP2* has not been described in bladder cancer, but was earlier identified as having pro-invasion activities in human melanoma cells

**Table 5** Univariate and multivariate Cox regression analysis of progression-free survival

	Univariate analysis		Multivariate analysis	
	Hazard ratio (95% CI)	P-value	Hazard ratio (95% CI)	P-value
Age (per 5 year increment)	1.21 (1.02–1.43)	<b>0.024</b>	1.02 (0.85–1.23)	0.806
Sex (men vs women)	0.62 (0.31–1.25)	0.184		
Stage (T1 vs Ta)	4.07 (2.13–7.77)	< <b>0.001</b>	1.64 (0.63–4.28)	0.311
Grade (high vs low + PUNLMP)	4.82 (2.25–10.34)	< <b>0.001</b>	1.25 (0.41–3.86)	0.703
Concomitant CIS <sup>a</sup> (presence vs absence)	2.37 (1.06–5.30)	<b>0.035</b>		
Treatment (BCG and MMC vs no treatment)	0.23 (0.09–0.58)	<b>0.002</b>	0.20 (0.07–0.57)	<b>0.003</b>
12-gene PCR signature (high risk vs low risk)	9.36 (4.66–18.81)	< <b>0.001</b>	8.30 (3.66–18.84)	< <b>0.001</b>

Abbreviations: BCG = bacillus Calmette-Guerin; CI = confidence interval; CIS = concurrent carcinoma *in situ*; MMC = mitomycin C. <sup>a</sup>CIS status was only known for 69 patients and consequently not included in the multivariate analysis. Bold indicates significant P-values (P < 0.05).

(Kabbarah *et al*, 2010). *NEK1* expression was identified in this study to be downregulated in progressing tumours compared with non-progressing tumours. Recently, this molecule has been shown to be involved in maintenance of chromosomal stability, and dysregulation of the molecule was associated with chromosomal instability (Chen *et al*, 2011). *COL4A3BP* expression has been found increased in drug-resistant cell lines, and it has been suggested to be a target for chemotherapy-resistant cancers (Swanton *et al*, 2007).

In conclusion, we have delineated a PCR-based 12-gene signature for disease progression in patients with non-muscle invasive bladder cancer. Based on molecular analysis of intra-patient variation, we suggest that this PCR-based test should be

applied to tumours from all patient visits in order to increase test sensitivity.

## ACKNOWLEDGEMENTS

The study was supported by The John and Birthe Meyer Foundation, the Danish Cancer Society, the Ministry of Technology and Science, and the Lundbeck Foundation. Furthermore, the research leading to these results has received funding from the European Community's Seventh Framework program FP7/2007-2011 under grant agreement no. 201663.

## REFERENCES

- Aaboe M, Marcussen N, Jensen KM, Thykjaer T, Dyrskjot L, Orntoft TF (2005) Gene expression profiling of noninvasive primary urothelial tumours using microarrays. *Br J Cancer* **93**(10): 1182–1190
- Billerey C, Chopin D, Aubriot-Lorton MH, Ricol D, Gil Diez de Medina S, Van Rhijn B, Bralet MP, Lefrere-Belda MA, Lahaye JB, Abbou CC, Bonaventure J, Zafrani ES, van der Kwast T, Thiery JP, Radvanyi F (2001) Frequent FGFR3 mutations in papillary non-invasive bladder (pTa) tumours. *Am J Pathol* **158**(6): 1955–1959
- Chen Y, Chen CF, Chiang HC, Pena M, Polci R, Wei RL, Edwards RA, Hansel DE, Chen PL, Riley DJ (2011) Mutation of NIMA-related kinase 1 (*NEK1*) leads to chromosome instability. *Mol Cancer* **10**(1): 5
- Dyrskjot L, Kruhoffer M, Thykjaer T, Marcussen N, Jensen JL, Moller K, Orntoft TF (2004) Gene expression in the urinary bladder: a common carcinoma in situ gene expression signature exists disregarding histopathological classification. *Cancer Res* **64**(11): 4040–4048
- Dyrskjot L, Thykjaer T, Kruhoffer M, Jensen JL, Marcussen N, Hamilton-Dutoit S, Wolf H, Orntoft TF (2003) Identifying distinct classes of bladder carcinoma using microarrays. *Nat Genet* **33**(1): 90–96
- Dyrskjot L, Zieger K, Kruhoffer M, Thykjaer T, Jensen JL, Primdahl H, Aziz N, Marcussen N, Moller K, Orntoft TF (2005) A molecular signature in superficial bladder carcinoma predicts clinical outcome. *Clin Cancer Res* **11**(11): 4029–4036
- Dyrskjot L, Zieger K, Orntoft TF (2007a) Recent advances in high-throughput molecular marker identification for superficial and invasive bladder cancers. *Front Biosci* **12**: 2063–2073
- Dyrskjot L, Zieger K, Real FX, Malats N, Carrato A, Hurst C, Kotwal S, Knowles M, Malmstrom PU, de la Torre M, Wester K, Allory Y, Vordos D, Caillaud A, Radvanyi F, Hein AM, Jensen JL, Jensen KM, Marcussen N, Orntoft TF (2007b) Gene expression signatures predict outcome in non-muscle-invasive bladder carcinoma: a multicenter validation study. *Clin Cancer Res* **13**(12): 3545–3551
- Ehdaie B, Theodorescu D (2008) Predicting tumor outcomes in urothelial bladder carcinoma: turning pathways into clinical biomarkers of prognosis. *Expert Rev Anticancer Ther* **8**(7): 1103–1110
- Hartmann A, Rosner U, Schlake G, Dietmaier W, Zaak D, Hofstaedter F, Knuechel R (2000) Clonality and genetic divergence in multifocal low-grade superficial urothelial carcinoma as determined by chromosome 9 and p53 deletion analysis. *Lab Invest* **80**(5): 709–718
- Hermann GG, Horn T, Steven K (1998) The influence of the level of lamina propria invasion and the prevalence of p53 nuclear accumulation on survival in stage T1 transitional cell bladder cancer. *J Urol* **159**(1): 91–94
- Jemal A, Siegel R, Xu J, Ward E (2010) Cancer statistics, 2010. *CA Cancer J Clin* **60**(5): 277–300
- Kabbarah O, Nogueira C, Feng B, Nazarian RM, Bosenberg M, Wu M, Scott KL, Kwong LN, Xiao Y, Cordon-Cardo C, Granter SR, Ramaswamy S, Golub T, Duncan LM, Wagner SN, Brennan C, Chin L (2010) Integrative genome comparison of primary and metastatic melanomas. *PLoS One* **5**(5): e10770
- Kompier LC, Lurkin I, van der Aa MN, van Rhijn BW, van der Kwast TH, Zwarthoff EC (2010) FGFR3, HRAS, KRAS, NRAS and PIK3CA mutations in bladder cancer and their potential as biomarkers for surveillance and therapy. *PLoS One* **5**(11): e13821
- Millan-Rodriguez F, Chechile-Toniolo G, Salvador-Bayarri J, Palou J, Algaba F, Vicente-Rodriguez J (2000) Primary superficial bladder cancer risk groups according to progression, mortality and recurrence. *J Urol* **164**(3 Pt 1): 680–684
- Ramaswamy S, Ross KN, Lander ES, Golub TR (2003) A molecular signature of metastasis in primary solid tumors. *Nat Genet* **33**(1): 49–54
- Rosenwald A, Wright G, Chan WC, Connors JM, Campo E, Fisher RI, Gascoyne RD, Muller-Hermelink HK, Smeland EB, Giltman JM, Hurt EM, Zhao H, Averett L, Yang L, Wilson WH, Jaffe ES, Simon R, Klausner RD, Powell J, Duffey PL, Longo DL, Greiner TC, Weisenburger DD, Sanger WG, Dave BJ, Lynch JC, Vose J, Armitage JO, Montserrat E, Lopez-Guillermo A, Grogan TM, Miller TP, LeBlanc M, Ott G, Kvaloy S, Delabie J, Holte H, Krajci P, Stokke T, Staudt LM (2002) The use of molecular profiling to predict survival after chemotherapy for diffuse large-B-cell lymphoma. *N Engl J Med* **346**(25): 1937–1947
- Rozen S, Skaletsky H (2000) Primer3 on the WWW for general users and for biologist programmers. *Methods Mol Biol* **132**: 365–386
- Swanton C, Marani M, Pardo O, Warne PH, Kelly G, Sahai E, Elustondo F, Chang J, Temple J, Ahmed AA, Brenton JD, Downward J, Nicke B (2007) Regulators of mitotic arrest and ceramide metabolism are determinants of sensitivity to paclitaxel and other chemotherapeutic drugs. *Cancer Cell* **11**(6): 498–512
- Sylvester RJ, van der Meijden AP, Oosterlinck W, Witjes JA, Bouffouix C, Denis L, Newling DW, Kurth K (2006) Predicting recurrence and progression in individual patients with stage Ta T1 bladder cancer using EORTC risk tables: a combined analysis of 2596 patients from seven EORTC trials. *Eur Urol* **49**(3): 465–466; discussion 475–477
- Thorsen K, Sorensen KD, Brems-Eskildsen AS, Modin C, Gaustadnes M, Hein AM, Kruhoffer M, Laurberg S, Borre M, Wang K, Brunak S, Kraimer AR, Topping N, Dyrskjot L, Andersen CL, Orntoft TF (2008) Alternative splicing in colon, bladder, and prostate cancer identified by exon-array analysis. *Mol Cell Proteomics* **7**: 1214–1224
- van't Veer LJ, Dai H, van de Vijver MJ, He YD, Hart AA, Mao M, Peterse HL, van der Kooy K, Marton MJ, Witteveen AT, Schreiber GJ, Kerkhoven RM, Roberts C, Linsley PS, Bernards R, Friend SH (2002) Gene expression profiling predicts clinical outcome of breast cancer. *Nature* **415**(6871): 530–536
- van de Vijver MJ, He YD, van't Veer LJ, Dai H, Hart AA, Voskuil DW, Schreiber GJ, Peterse JL, Roberts C, Marton MJ, Parrish M, Atsma D, Witteveen A, Glas A, Delahaye L, van der Velde T, Bartelink H, Rodenhuis S, Rutgers ET, Friend SH, Bernards R (2002) A gene-expression signature as a predictor of survival in breast cancer. *N Engl J Med* **347**(25): 1999–2009
- van Tilborg AA, de Vries A, de Bont M, Groenfeld LE, van der Kwast TH, Zwarthoff EC (2000) Molecular evolution of multiple recurrent cancers of the bladder. *Hum Mol Genet* **9**(20): 2973–2980
- Wang R, Morris DS, Tomlins SA, Lonigro RJ, Tsodikov A, Mehra R, Giordano TJ, Kunju LP, Lee CT, Weizer AZ, Chinnaiyan AM (2009) Development of a multiplex quantitative PCR signature to predict progression in non-muscle-invasive bladder cancer. *Cancer Res* **69**(9): 3810–3818
- Zieger K, Dyrskjot L, Wiuf C, Jensen JL, Andersen CL, Jensen KM, Orntoft TF (2005) Role of activating fibroblast growth factor receptor 3 mutations in the development of bladder tumours. *Clin Cancer Res* **11**(21): 7709–7719

This work is published under the standard license to publish agreement. After 12 months the work will become freely available and the license terms will switch to a Creative Commons Attribution-NonCommercial-Share Alike 3.0 Unported License.