Genomic imbalances in *Schistosoma*-associated and non-*Schistosoma*-associated bladder carcinoma. An array comparative genomic hybridization analysis

Gemma Armengola,b,*, Saad Eissa c, Juan José Lozando d, Sohair Shoman c, Lauro Sumoy d, María Rosa Caballín a, Sakari Knuutila b

a U. Antropologia, Departament de Biologia Animal, Biologia Vegetal i Ecologia, Faculty of Sciences, Universitat Autònoma de Barcelona, Bellaterra, Barcelona 08193, Spain
b Department of Pathology, Haartman Institute and HUSLAB, University of Helsinki and Helsinki University Central Hospital, Haartmaninkatu 3, 4th Floor, FI-00014, Helsinki, Finland
c Department of Pathology, National Cancer Institute, Kasr Al Ein St., Fom El Khalig, Cairo, Egypt
d Programa de Bioinformàtica i Genòmica, Centre de Regulació Genòmica (CRG), C/Dr. Aiguader, 88, 08003 Barcelona 08193, Spain

Received 7 February 2007; received in revised form 17 April 2007; accepted 17 April 2007

Abstract
Carcinoma of the urinary bladder is the most common malignancy in many tropical and subtropical countries due to endemic infection by *Schistosoma hematobium* (bilharzia). In the current study, we performed a high-resolution analysis of gene copy number amplifications using array comparative genomic hybridization to compare DNA copy number changes in pools of *Schistosoma*-associated (SA) and non-*Schistosoma*-associated (NSA) bladder cancer (BC). Many DNA copy number changes were detected in all studies, with multiple gains and losses of genetic material. The most frequent alterations were gains on 5p15.2~p15.33, 8q13.1, and 11q13, and losses on 8p21.3~p22 and 22q13. Even when SA pools showed no *Schistosoma*-specific gene copy number profiling as compared to NSA pools, some genes seemed to be gained (*ELN* on 7q11.23) and some lost (*PRKAG3* on 2q35 and *PRDM6* on 5q23.2) in SA-SCC. The following genes were gained in all histopathologic categories: *SRC* (20q11.23), *CEBPB* (20q13.13), and *GPR9* (Xq13.1). Our study did not provide clear evidence of differences in carcinogenesis of SA-BC and NSA-BC. © 2007 Elsevier Inc. All rights reserved.

1. Introduction
Carcinoma of the urinary bladder is the most common malignancy in many tropical and subtropical countries due to endemic infection by *Schistosoma hematobium* (bilharzia). *Schistosoma*-associated bladder cancer (SA-BC) is defined by characteristic pathology, and cellular and molecular biology that differ from non-*Schistosoma*-associated bladder cancer (NSA-BC) [1]. In Africa, SA-BC tends to present as squamous cell carcinoma (SCC) with invasive growth (pT2~T4) in relatively young individuals [2]. Few studies have analyzed the cytogenic and molecular genetic abnormalities in SA-BC [3~7], and to our knowledge, only one has compared DNA copy number changes in SA-BC and NSA-BC [8]. El Rifai et al. [8], Muscheck et al. [6], and Fadl-Elmula et al. [7] used metaphase comparative genomic hybridization (CGH) to obtain an overview of the chromosomal alterations in SA-BC. The current array CGH platform provides a number of advantages over the use of chromosomes, including higher-density region-specific coverage, direct mapping of aberrations to the genome sequence, and higher throughput [9]. Here, we performed a high-resolution analysis of gene copy number changes using array CGH to compare genomic imbalances in SA-BC and NSA-BC.

2. Materials and methods

2.1. Tumor samples and DNA
A total of 21 fresh-frozen bladder carcinoma samples were combined into six pools according to the
histopathologic type and Schistosoma infection status: two pools of SA-SCC (pool 1 with high-grade Bilharzia infection and pool 5 with low-grade Bilharzia infection); two pools of NSA-SCC (pools 2 and 6 without Bilharzia infection); one pool of SA-transitional cell carcinomas (TCC) (pool 3 with three cases of high-grade Bilharzia infection and one case of low-grade infection); and one pool of NSA-TCC (pool 4 without Bilharzia infection; Table 1). The tumors, obtained from the National Cancer Institute (Cairo, Egypt), were graded according to the World Health Organization classification [10].

Genomic DNA was extracted using the phenol-chloroform method. Before extraction, the pathologist (S.E.) determined that the proportion of tumor cells was higher than 50%. Reference DNA for CGH hybridizations was extracted from peripheral blood lymphocytes of four healthy individuals.

2.2. CGH on cDNA microarrays (array CGH)

Array CGH was performed on commercial cDNA microarrays (Human 1.0; Agilent Technologies; Palo Alto, CA, USA) as described earlier [11]. The array contains clones of about 13,000 genes throughout the genome. Of these clones, 982 had replicates on the array. Reference and sample DNA (20 μg) were digested with AluI and Rsal restriction enzymes (Life Technologies, Rockville, MD, USA). Digested samples were purified using the basic phenol-chloroform method. A total of 6 μg of digested tumor and reference DNA were then labeled with Cy3-dUTP and Cy5-dUTP (Amersham Biosciences, Piscataway, NJ, USA), respectively, using the RadPrime DNA labeling system (Gibco BRL, Gaithersburg, MD, USA). Cot-1 (150 μg), yeast tRNA (300 μg), poly dA (60 μg), and poly dT (60 μg) were added to the combined labeled DNA, and the mixture was concentrated to a total volume of 25 μL using microcon YM-30 filters (Millipore, Billerica, MA, USA) and adjusted to contain 3.4 × standard saline citrate (SSC) and 0.3% sodium dodecyl sulfate in the final volume. After denaturation (100°C for 1.5 minutes) and a 30-minute Cot-1 pre-annealing step (37°C), the probe was hybridized to the array under a glass coverslip at 65°C for 17 hours. After hybridization, the arrays were washed in 0.1% sodium dodecyl sulfate (3 minutes), 0.5x SSC/0.01% sodium dodecyl sulfate (5 minutes), and 0.06x SSC (3 minutes). Slides were centrifuged at 400 g for 2 minutes and scanned using Agilent’s laser confocal scanner (G2565AA).

Microarray images were analyzed using Agilent’s Feature Extraction software (version 6.1.1; Agilent Technologies) with the locally weighted linear-regression curve fit option. Log2 ratios were calculated from Cy3/Cy5 channels and further normalized. Measurements flagged as unreliable by Feature Extraction were excluded from subsequent analyses. Genomic alignment information was retrieved from the University of California at Santa Cruz’s Genome Browser database [12].

To analyze the results, individually affected areas were defined for each pool. We computed LOWESS smoothing lines in the data with an f span equal to 0.2 [13]. These parameters were determined by comparisons between the metaphase CGH and array CGH results obtained in our laboratory. To identify statistically significant transitions in the copy number, the variation of the copy number was defined as log2 ± 2 SD of the middle 50% quantile of data. Changes were considered reliable only when at least four different sequential clones were deleted or amplified.

In addition, we used a developed statistical segmentation algorithm termed circular binary segmentation (CBS), which parses the probe ratio data into segments of similar mean after taking variance into account. The algorithm analyzes one chromosome at a time and recursively identifies the best possible segmentation within that chromosome. Each proposed split is accepted or rejected on the basis of the probability that the difference in mean could have arisen by chance. This probability is determined by using a randomization method [14].

Furthermore, we performed a detailed analysis of single genes that were gained or lost in all BC pools and in all three SA-BC pools, in all two SA-SCC pools, or in the SA-TCC pool but not in their counterparts. Specific genes were considered to have copy number changes when both replicas (when available) had the same pattern (above or under the threshold values) and the neighboring probes followed a similar pattern.

<table>
<thead>
<tr>
<th>Pool</th>
<th>Sample code</th>
<th>Sex</th>
<th>Type</th>
<th>Grade</th>
<th>Stage</th>
<th>Bilharzia infection</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>9</td>
<td>M</td>
<td>SCC</td>
<td>G2</td>
<td>T3a</td>
<td>++</td>
</tr>
<tr>
<td>10</td>
<td>M SCC G2</td>
<td>T3b</td>
<td>++</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>M SCC G2</td>
<td>T3b</td>
<td>++</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>M SCC G2</td>
<td>T3b</td>
<td>++</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>11</td>
<td>M</td>
<td>SCC</td>
<td>G3</td>
<td>T2</td>
<td>-</td>
</tr>
<tr>
<td>21</td>
<td>F SCC G2</td>
<td>T3b</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>33</td>
<td>M SCC G2</td>
<td>T3b</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>38</td>
<td>M SCC G2</td>
<td>T3b</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>F</td>
<td>TCC</td>
<td>G3</td>
<td>T3b</td>
<td>++</td>
</tr>
<tr>
<td>4</td>
<td>M TCC G2</td>
<td>T3b</td>
<td>++</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>M TCC G3</td>
<td>T3b</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>M TCC G2</td>
<td>T3b</td>
<td>++</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>6</td>
<td>M</td>
<td>TCC</td>
<td>G2</td>
<td>T3b</td>
<td>-</td>
</tr>
<tr>
<td>8</td>
<td>M TCC G2</td>
<td>T3b</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>41</td>
<td>M TCC G2</td>
<td>T3b</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>22</td>
<td>M</td>
<td>SCC</td>
<td>G3</td>
<td>T3b</td>
<td>+</td>
</tr>
<tr>
<td>25</td>
<td>M SCC G2</td>
<td>T2</td>
<td>+</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>3</td>
<td>F</td>
<td>SCC</td>
<td>G3</td>
<td>T3b</td>
<td>-</td>
</tr>
<tr>
<td>28</td>
<td>M SCC G3</td>
<td>T3b</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>M SCC G3</td>
<td>T3b</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>39</td>
<td>F SCC G3</td>
<td>T3b</td>
<td>-</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: SCC, squamous cell carcinoma; TCC, transitional cell carcinoma; ++, high-grade Bilharzia infection; +, low-grade Bilharzia infection.
3. Results and discussion

We pooled DNA from SA-SCC and SA-TCC tumors as well as NSC-SCC and NSA-TCC tumors. The value of pooled DNA has been shown to be advantageous in detecting recurrent changes associated with specific histopathologic or clinical features [15]. Pooled DNA reveal recurrent primary changes covering secondary changes that vary from case to case. Other advantages of pooled genomic samples are that small DNA amounts permit the analysis, thereby reducing the need for hybridization procedures.

Many DNA copy number changes were detected in all experiments, with multiple gains and losses of genetic material (Table 2). The most frequent alterations were gains on 5p15.2–p15.33, 8q13.1, and 11q13, and losses on 8p21.3–p22 and 22q13. The original microarray data are deposited in the CanGEM database (www.cangem.org).

Fig. 1 shows an example of whole-genome array CGH profile from pool 3. No differences were found in copy number profiling between SA-BC and NSA-BC; i.e., the pools of SA tumors showed no schistosomiasis-specific changes compared to pools of NSA tumors. The comparison between SA-TCC and NSA-TCC pools gave similar results, as did the comparison between SA-SCC and NSA-SCC pools. According to metaphase CGH studies of SA-BC, all detected imbalances have been repeatedly reported in NSA-BC, suggesting that the cytogenetic profiles of chemical- and bilharzia-induced carcinomas are largely similar [6–8]. Moreover, our results of TCC showed alterations that have previously been reported in other array-based CGH studies on NSA-TCC [16–18].

Further analyses of individually studied genes did not show differences in copy number changes between SA-BC and NSA-BC either. However, some differences were found between SA-SCC and NSA-SCC. Notably, the elastin (ELN) gene (7q11.23) was gained in SA-SCC pools but not in NSA-SCC pools. Increased ELN expression has been observed in human breast cancer cell lines [19]. Furthermore, the AMP-activated protein kinase (AMPK)-gamma noncatalytic subunit (PRKAG3) and PR domain containing protein 6 (PRDM6) genes (on 2q35 and 5q23.2, respectively) were lost in SA-SCC but not in NSA-SCC. Recent studies have suggested a link between AMPK and cancer; a number of molecules, including mTOR and fatty acid synthase (FAS), which are necessary for the growth of many tumors, are inhibited by AMPK [20]. PRDM6 is a PR domain gene (PRDM), and other genes in this family are considered tumor suppressors because they are silenced or inactivated in several tumors [21–23].

Moreover, the detailed analysis of individual genes showed a set of genes with the same copy number changes in all bladder carcinoma pools. The following interesting genes were gained: v-src sarcoma (Schmidt-Ruppin A-2) viral oncogene homolog (avian) (SRC on 20q11.23), CCATT/enhancer-binding protein (C/EBP)-beta (CEBPB on 20q13.13), and G protein–coupled receptor 9 (GPR9 on Xq13.1). SRC is a known proto-oncogene that has been observed both overexpressed and highly activated in a number

Table 2
DNA copy number changes detected by array CGH for each pool of bladder cancer

<table>
<thead>
<tr>
<th>Pools</th>
<th>Gains</th>
<th>Losses</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5p15.33 (0.9), 12q14.1-q14.2 (2.1), 22q11.21 (0.6)</td>
<td>4p15.32-p15.31 (4.9), 7q34-q35 (0.9), 8p21.3-p22 (3.3), 22q13.1 (0.3)</td>
</tr>
<tr>
<td>2</td>
<td>7p13 (0.3), 7p14.3-p15.1 (2.3), 7p15.2 (0.06), 8q13.1 (0.8), 11q13.1 (0.07 and 0.2), 11q13.3 (0.7), 14q23.31-q23.32 (1.3)</td>
<td>—</td>
</tr>
<tr>
<td>3</td>
<td>10p15.1 (1.4), 10p14.15.1 (1.5)</td>
<td>—</td>
</tr>
<tr>
<td>4</td>
<td>5p15.2-p15.33 (11), 17q11.2 (0.7 and 0.1), 18p11.22 (0.5), 19q13.12 (0.1), 19q13.2 (1.5 and 0.5), 20q11.21 (0.4), 20q13.12 (0.1), 20q13.13 (0.6), 20q13.32-q13.33 (2.6), 20q13.33 (0.8)</td>
<td>6q13 (3.3), 14q13.2 (0.7), 14q24.2 (2.4), 15q14 (4.2), 15q15.3 (0.3), 17p12-p13.1 (4.6), 18q21.2-q21.31 (1.9), 22q13.1 (0.3)</td>
</tr>
<tr>
<td>5</td>
<td>1q23.3 (1.4), 8q12.3-13.2 (4), 8q23.2-q24.12 (9.6), 9p24.3 (2.1), 11q13.4 (0.6), 16q12.1 (1.4)</td>
<td>3p21.31 (0.6), 3p11.2-q11.2 (12), 3q13.13-q13.3 (3.8), 5q14.1-q14.3 (6.7), 5q23.1-q23.3 (8.7), 5q31.2 (0.1), 5q33.1 (0.8), 8p21.3-p22.3 (2.5), 8p11.21 (0.4), 9q24.1 (1.5), 9q21.1-p21.3 (10.4), 11q22.2-q22.3 (2), 11q24.3 (1)</td>
</tr>
<tr>
<td>6</td>
<td>—</td>
<td>15q25.3 (0.9)</td>
</tr>
</tbody>
</table>

Sizes in parentheses are in mega–base pairs.
of human cancers. It has been found to be a critical component of multiple signaling pathways that regulate proliferation, survival, metastasis, and angiogenesis [24]. In addition, expression of CEBPB mRNA and protein is markedly increased in several tumors [25,26], and G protein-coupled receptors are expressed at a significantly higher level in some cancer tissues and may play a role in cancer progression [27].

Our array CGH analysis did not indicate differences in gene copy profiling between SA-BC and NSA-BC. However, we found an elevated copy number of the ELN gene and decreased copy numbers of PRKAG3 and PRDM6 in SA-SCC bladder tumors. Our results call for further studies on more sensitive array CGH platforms and/or by fluorescence in situ hybridization with gene-specific probes to confirm the findings. Further analyses using reverse-transcription polymerase chain reaction or Western blotting are also warranted to determine the expression status of candidate genes.

Acknowledgments

The authors thank Tiina Wirtanen for skillful help in hybridizations.

References