Paired Exome Analysis Reveals Clonal Evolution and Potential Therapeutic Targets in Urothelial Carcinoma

Philippine Lamy1, Iver Nordentoft1, Karin Birkenkamp-Demtroeder1, Mathilde Borg Houlberg Thomsen1, Palle Villesen2, Søren Vang1, Jakob Hedegaard1, Michael Borre3, Jørgen Bjerggaard Jensen3, Søren Høyer4, Jakob Skou Pedersen1, Torben F. Ørntoft1, and Lars Dyrskjøt1

Abstract

Greater knowledge concerning tumor heterogeneity and clonality is needed to determine the impact of targeted treatment in the setting of bladder cancer. In this study, we performed whole-exome, transcriptome, and deep-focused sequencing of metachronous tumors from 29 patients initially diagnosed with early-stage bladder tumors (14 with nonprogressive disease and 15 with progressive disease). Tumors from patients with progressive disease showed a higher variance of the intrapatient mutational spectrum and a higher frequency of APOBEC-related mutations.

Introduction

Bladder cancer is a major cause of mortality worldwide (1). More than 90% of bladder tumors are classified as urothelial carcinoma, and around 25% of patients are initially diagnosed with muscle-invasive bladder cancer (MIBC). The remaining 75% are diagnosed with non-muscle invasive bladder cancer (NMIBC), with high recurrence rates and moderate progression rates (~15%) to MIBC (1, 2). Identifying clinically informative alterations in bladder cancer may help stratifying patients to follow-up regimens, chemotherapy, and targeted treatment, which may ultimately increase survival. However, knowledge on tumor heterogeneity and clonality of potential therapeutic targets is needed to administer optimal treatment.

Bladder cancer has a heterogeneous clinical presentation, and recurrent tumors are often observed at shifting locations in the bladder. Furthermore, patients with multiple tumors are at higher risk of recurrence, progression, and cancer-related death (1). Multifocal tumors are hypothesized to originate from a field cancerization of the bladder, where tumors develop from pre-neoplastic fields of transformed urothelial cells (3). The field cancerization model is supported by frequently observed genetic instability and mutations in normal-appearing urothelium (4) and dysplasia or carcinoma in situ (CIS) in the surrounding urothelium (5). In addition, analysis of gene expression in tumors and adjacent normal-appearing urothelium has shown expression patterns in the normal urothelium similar to those observed in the tumor lesions (5).

The recent identification of mutational signatures characteristic for each cancer type has provided insight into the mutational heterogeneity and mutational mechanisms as well as repair deficiencies that lead to a given cancer (6). The continued acquisition of mutations may explain the shift from benign to malignant tumors over time and could in theory guide preventive measures. The mutational landscape in NMIBC has recently been described (7–9). Furthermore, the APOBEC enzyme family that creates specific mutational signatures in single-stranded DNA was shown to be prevalent in bladder cancer, and it may fuel the oncogenic transformation (10, 11). Finally, deep sequencing has recently been used to identify the clonal composition of driver alterations in several cancers (7, 12, 13).

For optimal targeted treatment, it is important to define the ancestral mutations, as these may be ideal therapeutic targets (present in all cancer cells). Here, we defined the mutational events occurring over time in multiple paired bladder tumors to identify phylogenetic relations between metachronous tumors. A common origin was found for all tumor pairs, but driver mutations in known tumor suppressors and oncogenes were not predominantly found to be early events. Potential therapeutic
targets were identified as both ancestral (shared) and tumor-specific alterations. We found a higher intrapatient variation of the tumor mutation spectrum and frequent APOBEC-related mutation signatures in patients with progressive disease (PD). In addition, we found high temporal and low spatial mutational heterogeneity, suggesting that targeted treatment decisions should ideally be based on analysis of biopsies from several tumors.

**Materials and Methods**

**Sample preparation**

This study includes data from a previous study in which metachronous tumors from 4 patients were analyzed together with single tumors from 20 patients (8). In total, 75 UCC samples from 38 patients were obtained fresh from surgery, embedded in Tissue-Tek, and stored at –80°C. Two or more metachronous tumors were included for 29 patients. Inclusion criteria are defined in the Supplementary Data. In addition, for one patient (p39), DNA was extracted from eight laser microdissected subregions from a single tumor. Leucocyte DNA from initial visits was sequenced as reference. Genomic DNA and total RNA were extracted from serial cryosections. All patients gave their written informed consent, and the study was approved by the National Committees on Health Research Ethics (#1300174).

**NGS library construction and sequencing**

Libraries for whole-exome sequencing (WES) were made using the Illumina TruSeq DNA Kit and NimbleGen SeqCap EZ v3.0. Whole-transcriptome RNA sequencing (RNA-seq) was prepared from rRNA-depleted total RNA using the Ribo-Zero and ScriptSeq (Epicentre). WES for patient p39 was made using the Illumina Nextera Rapid Exome Capture Expanded Kit. Libraries for ultradeep targeted resequencing were constructed using the Illumina TruSeq Custom Amplicon v1.5 Kit and contained 1,530 amplicons. Amplicons were designed using DesignStudio software v1.5 (Illumina). Sequencing was performed on the Illumina HiSeq2000 and NextSeq 500 platforms.

**Sequence processing**

DNA samples were mapped to hg19, marked for duplicates, and the alignments were recalibrated and realigned using the picard (http://picard.sourceforge.net/) and GATK suites (14). SNPs were called using GATK HaplotypeCaller and somatic mutations (SNV) and indels were called using MuTect (15) and the Somatic Indel Detector from GATK. Indels shared by multiple patients were excluded. Functional annotation of mutations were performed using Snpeff (16), and SNVs were classified in two categories reflecting the quality of the calls (Cat 1, high quality; Cat 2, low quality) and four tiers reflecting the impact of the call (Tier 0, nonsense or splice site mutations; Tier 1, missense mutations; Tier 2, silent mutations; Tier 3, noncoding mutations). RNA-seq data were mapped and quantified using the Tuxedo Suite (17).

**Mutational signatures**

For all samples, the sequence context (3’ and 5’ nucleotides) was extracted for all somatic mutations (Cat 1–2 Tier 0–3), leading to 96 classes of mutations. Each sample mutational profile was projected onto 22 known signatures (18), and only signatures contributing to more than 7% of the mutations were retained. This was done iteratively until no more signatures were removed. The six most frequent signatures were used, and their contributions were recalculated.

**Allele-specific expression**

We analyzed DNA and RNA sequencing data combined to investigate allele-specific expression (ASE) in the tumor samples. ASE was measured as the absolute difference in alternative allele frequency between RNA and DNA.

**Phylogenetic tree analysis**

For all patients with multiple samples, we investigated whether a mutation was present (minimum two reads). This was done for all mutations called as Cat 1 in any of the samples from a given patient. We then used the status of each mutation in all samples to find the most probable phylogenetic tree (see Supplementary Data for details). Confidence for internal nodes was estimated using bootstrap analysis.

**Copy number estimation**

The sequenza package was used to define the copy number variation from WES data (19).

**Cancer cell fraction calculation**

Cancer cell fraction (CCF) of Cat 1 Tier 0–1 mutations was estimated by integrating copy number estimates and normal cell contamination (histology) with the variant allele frequency as described in ref. 20. For each autosomal mutation with at least 30 reads, we obtained an estimate of the CCF and a 95% confidence interval (CI). We defined mutations as being clonal if the 95% CI included 0.95 and subclonal otherwise.

**Statistical inference of tumor cell populations using deep exome sequencing**

PyClone 0.12.7 (12) was used to infer subclonal populations.

**Statistical analysis**

t test statistics was used unless otherwise indicated.

**Data access**

The sequencing data reported were deposited to the European Genome-Phenome Archive (EGA; www.ebi.ac.uk/ega/) under accession nr. EGAS00001001686.

**Results**

**DNA and RNA sequencing**

WES was performed on 75 bladder tumors from 38 patients (Supplementary Table S1A and S1B). We analyzed two to five metachronous tumors from 29 patients and a single tumor from 9 patients. In total, 17 patients (35 tumors) had nonprogressive disease (NPD, patients with recurrent disease and no stage progression) and 21 patients (40 tumors) had PD (patients with transitions from Ta→T1→4 or from T1→T2→4). Tumor and matched germline DNA was sequenced to an average depth of 90× (28-450); Fig. 1A; Supplementary Table S1C). We identified an average of 711 (113-3331) SNVs per sample within the exome target regions, categorized by strength of evidence and expected functional impact (Snpeff; ref. 16) as described previously (Supplementary Table S2; ref. 8) as well as an average of 26 (5-104) indels per sample (Supplementary Data).
Most frequently hit genes by mutations (Cat 1 Tier 0 + 1 deleterious) and indels
Table S2B). Of these, an average of 110 (5-1,259) SNVs (Cat 1 Tier 0 + 1) and 8 (1-42) indels per tumor had a predicted effect on protein function (Fig. 1B). The number of mutations was not correlated to sequencing depth. Whole-transcriptome sequencing was performed for 71 of the 75 tumors. Samples were sequenced to a mean of 119.6 (7.8–234.6) million mapped reads (Fig. 1A; Supplementary Table S1D). Clinical data for all patients are summarized in Fig. 1C.

**Mutational spectrum of tumors from NPD and PD patients**

To evaluate the mutational landscape in bladder cancer, we identified the most frequently mutated genes (Supplementary Fig. S1). A subset of genes with mutations predicted as deleterious (SIFT, PROVEAN, and PolyPhen2) is presented in Fig. 1D (Tier 2). The most frequently mutated genes in the total patient cohort were **KDM6A** (53%), **ARID1A** (32%), **FGFR3** (26%), **TP53** (24%), **FAT1** (18%), **EP300** (16%), **RB1** (16%), and **STAG2** (16%; refs. 8, 9, 23). Compared with studies of MIBC, we observed higher mutation rates of **KDM6A**, **ARID1A**, **FGFR3**, **FAT1**, **RB1M01**, and a lower mutation rate of **TP53** (Fig. 1D and Supplementary Fig. S1; refs. 23–26). There was no significant difference in the mutation rates when comparing the NPD and PD patients (10.3 vs. 14.6 mutations/Mb; t test with unequal variances, P = 0.20, SNVs Cat 1–2 + indels). Furthermore, no significant difference was observed when considering only the first Ta tumor from each patient in each group (8.8 vs. 12.1 mutations/Mb; t test with unequal variances, P = 0.37, SNVs Cat 1–2 + indels). The variation in the PD group was significantly higher for both comparisons (Bartlett test, P = 0.001 and P = 0.004).

We observed higher intratumor variation of the mutational spectrum for metachronous tumors from patients with PD compared with tumors from patients with NPD (χ² test, P = 0.0013; Fig. 2A), pointing toward additional mutational mechanisms in the PD group, or more variation in repair capacity. Using RNA-seq data, we classified all samples according to a previously reported CIS signature (5) using ConsensusClusterPlus (Fig. 1C; ref. 27). The CIS signature is associated with high-risk NMIBC (concomitant CIS) and disease progression (28). The CIS signature showed a significant correlation to progression in our study (χ² test; P < 0.02). When stratifying the mutational analysis based on samples with similar gene expression signatures, we did not observe a significant difference in SNVs or indels between the groups (Fig. 1D).

Analyzing the proportions of the individual mutation types (Cat 1) in the initial versus the most recent tumor showed a significant decrease (P < 0.017) of **C**→**T** mutations in the PD patients (Supplementary Fig. S2A–S2D). When comparing the change in proportion (mean of absolute changes) of the individual mutation types from the initial to the most recent tumor, we found a larger change in **C**→**G** and **C**→**T** (P < 0.05) in the PD patients compared with the NPD patients (Supplementary Fig. S2E). Taking the direction into account, only the proportion of **C**→**T** mutations remained significant (P < 0.01; Supplementary Fig. S2F).

To gain insight into the mutational processes and signatures, we included the 5′ and 3′ nucleotides and grouped the mutations (Cat 1 + 2) by context (Supplementary Fig. S3; ref. 6). The frequency of the 22 signatures described by Alexandrov and colleagues (6) was estimated in our 75 tumors (Fig. 2B). In most tumors, we observed several signatures (on average, 3.72 signatures). The six most abundant signatures (1A, 1B, 2, 6, 12, and 13) were used to define the mutational processes in the samples (Fig. 2C). Signature 1A was found with a higher frequency in the NPD group compared with the PD group (P < 0.0001), whereas no differences between groups were observed for the remaining 5 signatures (Supplementary Fig. S4). When focusing only on the first Ta tumors from the two groups, signature 1A was found with a higher frequency in the NPD group compared with the PD group (Mann-Whitney test, P < 0.0011), and signatures 1B and 6 were more frequent in the PD group (Mann-Whitney test, P = 0.0270 and P = 0.0295, respectively; Fig. 2D). Signatures 1A and 1B (characterized by **C**→**T** substitutions at Np CpG trinucleotides) are related to age (5-methyl-cytosine deamination). However, there was no age difference between patients at the time of the first Ta tumors in the NPD and PD groups (Mann-Whitney test, P = 0.12, median age 66 [39–84] vs. 67 [58–80]).

When comparing the initial versus the most recent tumor in the two groups, signature 1A was found more frequently in the progressed tumors in the PD group (Wilcoxon signed-rank test, P = 0.0048), whereas no difference was observed in the NPD group (Supplementary Fig. S5). The finding that signature 6 (C→T at NpCpG) is present in 60% of the tumors and is more abundant in the PD group is interesting, as this signature is a hallmark of DNA mismatch repair deficiency, not commonly described in bladder cancer.

Samples were classified according to the APOBEC signature, where “high APOBEC” was defined as a contribution of signatures 2 and 13 above 30% and low otherwise. About two thirds of the PD patients were classified as “high APOBEC” in at least one of their tumor samples compared with one third in NPD patients (Fig. 2E; 66% vs. 35%; Fisher’s exact test, P = 0.04). Interestingly, intratumor shift in APOBEC classification was observed in only one NPD patient (1/14), while it was seen in 53% (8/15) of PD patients (Fig 2E; Fisher’s exact test, P = 0.009). In PD patients with a classification shift, the “high APOBEC” classification was observed in the initial tumor for half of the patients. Consequently, the data indicate that APOBEC mutational activity is a late event, often tumor specific.
No correlation between the APOBEC score and the mutation rate (Cat 1 + 2 SNVs/Mb) was detected (Supplementary Fig. S6). Previous studies have shown a correlation between the APOBEC signature and APOBEC expression, especially APOBEC3B (10, 29). Here, transcript expression values of nine APOBEC genes did not show a correlation between expression and APOBEC signatures 2 + 13. Comparing APOBEC expression in Ta tumors from NPD versus PD patients showed no significant difference; however, APOBEC3A expression was borderline significantly higher in the PD group (Mann–Whitney test, \( P = 0.065 \); Supplementary Fig. S7). In conclusion, we identified no correlation between mutational load and expression of APOBEC transcripts. The reason for this could be that mRNA abundance in tumors does not represent abundance at the time of mutagenesis.

Recently, Chan and colleagues (30) showed that APOBEC3A and APOBEC3B can be further separated from the TCA motif based on the nucleotide 5′ of the consensus trinucleotide, where APOBEC3A (A3A) prefers a pyrimidine base and APOBEC3B (A3B) a purine base. Conducting this analysis showed that the signature of A3A contributed twice as much as A3B (Fig. 2F). Interestingly, we found the opposite pattern for the mRNA levels of A3A and A3B (Fig. 2G), in concert with the A3A enzyme being more mutagenic than A3B. We observed no differences in A3A and A3B across groups.

ASE

We performed a combined analysis of DNA and RNA to investigate ASE. First, we compared our tumor samples with samples from normal bladder urothelium from TCGA (31). Overall, cancer samples showed a higher ASE compared with normal samples [Wilcoxon rank sum test with continuity correction (WRST), \( P = 1.14e–96 \)], especially for tumor suppressor genes (WRST, \( P = 0.00169 \)) but also more broadly (‘other genes’: WRST, \( P = 3.89e–100 \); Supplementary Table S3A). Five genes, PLEC, ACTG2, SVIL, COL5A1, and MMP2, were found to have a higher ASE in bladder cancer than in normal bladder with an FDR below 0.25 (Supplementary Table S3B). We also observed that genes that were hit by a mutation had a significantly higher ASE (mean = 0.183) than nonmutated genes (mean = 0.165; WRST, \( P = 2.37e–07 \)). Interestingly, the ASE was directed toward the alternate allele for Tier 0 mutations (exact binomial test, \( P < 0.039 \)), while it was directed toward the alternate allele for all other tiers (exact binomial test, Tier 1: \( P < 8.8e–12 \); Tier 2: \( P < 2.6e–5 \)). Next, we compared the NPD and the PD groups and found higher ASE in the initial Ta tumors from PD patients compared with those from NPD patients (WRST, \( P = 1.18e–56 \)), with tumor suppressor genes showing a larger ASE than oncogenes (WRST, \( P = 0.0164 \) vs. \( P = 0.417 \); Supplementary Table S3C and S3D).

**Group and pairwise comparison of frequently mutated genes**

Comparison of the initial Ta tumors from each group identified possible early drivers of progression (LRP1B, VR3, PCDHA12, TLR1, RP111, BXX, ZNF717, ADAMTS18, TENM1, and SSPO) only mutated in Ta tumors from the PD group (Supplementary Fig. S8A). A comparison of the initial Ta tumors from NPD patients to the T1 or T2 tumors from PD patients (most recent tumor) is found in Supplementary Fig. S8B. A comparison of paired Ta tumors from NPD patients revealed frequent mutations in PIK3CA, RBM10, KDM6A, RANBP2, and FGFR3 were early events present in the initial Ta, whereas mutations in MKI67, CCDC168, MIA3, and OPN1LW primarily occurred late in tumorigenesis of recurrent Ta tumors (Supplementary Fig. S8C). Finally, comparing metachronous tumors from PD patients showed frequent mutations in genes like RYR3, RP111, KDM6A, FGFR3, PIK3CA, and ZNF717 in the initial tumor (Supplementary Fig. S8D). Interestingly, only 35% (7/20) of the mutations in these genes seemed to be retained in the progressed samples. However, visual inspection resulted in detection of additional recurrent low frequency SNVs (marked by orange in Supplementary Fig. S8D) in early (\( n = 7 \)) and late PD samples (\( n = 8 \)), increasing the retained mutations to 50% (10/20) in these genes.

**Temporal analysis of mutations in individual patients**

To gain further insight into the accumulation of genetic alterations, we constructed phylogenetic trees of disease evolution taking all mutations in Cat 1 Tier 0–2 into account. We hypothesized that mutations found in the ancestral branch, that is, observed in all samples from the same patient, probably characterized the field disease. We observed a tendency for longer ancestral branches [60.8% (31%–85%) vs. 46.5% (14%–87%) of the SNVs; \( P = 0.06 \) for the NPD patients (Fig. 3 and Supplementary Fig. S9)]. Furthermore, in 4 of 6 patients with three or more tumors analyzed, we observed that new mutations are only found in the latest tumors (Fig. 4 and Supplementary Fig. S10; Supplementary Table S4).

Next, we hypothesized that a larger fraction of mutations in the ancestral branch is clonal (present in all carcinoma cells) compared with the sample-specific branches, as they arose early (20). We calculated the CCF for all Cat 1 Tier 0–1 mutations in all samples as in ref. 20. Estimations were made only for mutations on autosomal chromosomes and with enough coverage (>30 reads). The clonal status is shown in Figs. 3 and 4 for all driver genes in patients with paired samples. When looking at patients with two tumors analyzed, we observed as expected a higher proportion of clonal mutations in the ancestral branch versus sample-specific branches (Wilcoxon signed-rank test one-tailed, \( P = 0.002 \); Supplementary Figs S11–S14).
Spatial heterogeneity
The analysis of metachronous tumors showed a high degree of intrapatient mutational heterogeneity. We therefore asked whether this reflected high intratumor heterogeneity, potentially related to tissue sampling bias, or whether it mainly reflected changes between metachronous tumors. To evaluate the degree of intratumor heterogeneity, we performed multiregional exome sequencing. Eight regions from a single muscle-invasive tumor biopsy were laser microdissected and exome sequenced to a mean depth of $61.1 \times (31.4–73.6)$ per sample (Fig. 5A). The regions were evolutionarily similar (Fig. 5B and C), with relatively few private mutations. However, allele frequencies of the SNVs in four adjacent regions were consistently higher compared with the remaining four regions (Fig. 5B, dark red), indicating a higher normal tissue contribution in these regions. A pool of sequences from all eight regions was analyzed with PyClone. Two main subclones were observed reflected by the two main clusters (Fig. 5D), underlining a low degree of heterogeneity. However, no biclonal pattern could be observed in the heatmap from the eight regions (Fig. 5B and C), pointing toward a similar intermix of the two subclones in all regions. This suggests that the heterogeneity observed at the bulk level is also present at the local level. Consequently, a minor fraction of the intrapatient variation appears to be due to intratumor heterogeneity.

Subclonal selection
To understand patterns of clonal selection in patients with metachronous tumors, we performed ultradeep sequencing of selected SNVs (30–105 for each patient) in 20 patients with two tumor samples ($n = 15$), three tumors ($n = 4$) and five tumors ($n = 1$). For all 1,800 SNVs, we achieved a median coverage of $6,809 \times (2,650–87,455)$ per sample (Supplementary Table S5). The clonal subpopulations were inferred using PyClone (12) with copy numbers derived from exome sequencing data (19). Overall, we observed relatively few subclones; in most cases, we observed an ancestral clone (defined as a set of mutations) present in all the cells from all tumors and one or two private clones in each of the tumors. The mutations from the private clones were either present in all cells (e.g., patient p09) or only in a subset of the cells (e.g., patient p02; Supplementary Figs. S15–S18).

Next, we combined subpopulation analyses with the mutational signature results. In patient p26, the phylogenetic tree showed that the initial Ta tumor deviated from the two most recent tumors. The mutational profile supported this, as we observed a strong APOBEC signature only in the two most recent tumors (Fig. 6A). Furthermore, the expression of APOBEC3B was approximately two times higher in the later tumors compared with the initial (Fig. 4A and Supplementary Fig. S19). Interestingly, the PyClone analysis showed the presence of a main subclone present in half of the cells and only found in the two most recent tumors (Fig. 6B). This could be interpreted as the development of a new clone, after the resection of the initial tumor, characterized by a strong APOBEC signature. Patient p19 showed a different pattern. Both the APOBEC signature and the PyClone analysis suggested that the most recent tumor was different from the first two. The T1 tumor was characterized by a strong APOBEC signature not seen in the two previously sampled tumors (Fig. 6C), and the expression of APOBEC3B was increased compared with the Ta tumors (Fig. 4A and Supplementary Fig. S19). Furthermore, the PyClone analysis suggested the development of a new subclone in the T1 tumor present in about half of the tumor cells (Fig. 6D).

Finally, we looked for potential therapeutic targets in the NPD and the PD group. We searched the DGIdb (The Drug Gene Interaction database; ref. 32). TARGET (tumor alterations relevant for genomics-driven therapy) available from the Broad Institute as well as the IntOGen database and focused our analysis on FDA-approved drugs. In total, we identified 19 altered potential therapeutic target genes, and the majority of the targets were found in the ancestral branches (Figs. 3 and 4; Supplementary Table S6). At least one potential therapeutic target in the ancestral branch was observed in 72% (21/29) of patients. Whether the proposed drugs are applicable to the different variants identified is beyond the scope of this analysis. Activating FGFR3 mutations were observed in 11 patients (S249C, Y373C, and the less common G370C mutation), and in PIK3CA, eight mutations were found, of which five are known to be activating mutations (E542K, E545K, Q546E, and H1047L). In total, 11 activating ancestral FGFR3 and PIK3CA mutations across 9 patients [31%, (9/29)] were identified, indicating that targeted treatments of these genes and pathways may be highly relevant in NMIBC.

Discussion
Detailed knowledge on tumor heterogeneity and early clonal mutations with disease-driving potential is needed to implement targeted treatment. Bladder cancer has one of the highest mutation rates, which entail increased possibility of mutational heterogeneity and increased likelihood of identifying potential therapeutic targets (6, 31). Here, we characterized genomic alterations in paired tumors from 29 patients initially diagnosed with NMIBC. We found a higher intrapatient variation of the tumor mutation spectrum in patients with PD, suggesting that new mutational processes are acquired during progression. In all patients, we observed shared ancestral mutations representing the field disease, documenting the clonal nature of bladder cancer as previously observed (8).

Tumor evolution has been assessed previously in bladder cancer, showing a clonal origin of metachronous tumors (8, 33, 34). Here, we observed a higher genomic diversity in the tumors from PD patients, consistent with the change in cellular properties acquired during disease progression. This diversity might not be explained only by a tissue sampling procedure, as we did not observe large differences in the mutational profile from eight regions within a single tumor biopsy. The observation needs...
to be confirmed in additional samples. Inferring tumor subpopulations by integrating deep sequencing data and copy number information can be done for a single sample, but it is more informative using multiple metachronous samples, where subclones can be traced over time (12). Single-cell sequencing performed on breast cancer xenografts has validated the technical procedure, as the main subclones found by PyClone analysis of bulk tumor biopsies were confirmed by the clonal genotypes from single cells (7, 12). Few subpopulations were inferred in triple-negative breast cancer (35) and in tumors from a few patients with bladder cancer (8, 36). In this study, we also observed few subclones and no statistical difference between patients with NPD and PD. However, our analysis may underestimate the number of clones, as it was conducted on a selected set of SNVs. To compensate for this, we also investigated the clonal status of the mutations, as described earlier (20, 35). This approach requires reliable estimations of the tumor cellularity and the copy number. Here, we showed a general higher proportion of clonal mutation in the ancestral branch that may characterize the field disease.

Figure 4.
Disease evolution in patients with three or more tumors analyzed. A, heatmaps of alterations in samples from patients having more than two metachronous samples. Red boxes designate mutations and intensity reflects their frequencies, black boxes indicate no mutation, and gray boxes no classification. Known tumor suppressors (T), oncogenes (O), and bladder intOGen genes are highlighted (red box, Tier 0; orange, Tier 1). *, genes potentially actionable using FDA-approved drugs. The APOBEC3B expression is shown below each heatmap. Heatmaps showing all gene names, allele frequencies, and gene expression levels are displayed in Supplementary Fig. S10 and Supplementary Table S4. B, phylogenetic trees (top, NPD patients; bottom, PD patients). The total number of mutations and the percentage of SNVs belonging to each branch are indicated. Long ancestral branches were shortened for graphical purpose (two crossing lines). Tumor suppressor genes, oncogenes, and intOGen genes with Tier 0–1 SNVs are highlighted (ancestral, black font; sample specific, colored fonts; intern branches, gray fonts) together with their clonal status as defined in ref. 20 (red, clonal; blue, subclonal; and gray, not defined). *, genes potentially actionable using FDA-approved drugs. Time scale, time between the tumor removals. Bootstrap P values are displayed to the left of the trees.

Figure 5.
Spatial heterogeneity in bladder cancer. A, WES was applied to DNA from eight regions of a tumor from patient P39. B, heatmap of alterations in the eight regions. Bottom, boxplot of the distribution of allele frequencies. C, top, the regions were clustered (Spendall t correlation) using frequencies for all SNVs found in the tumor; bottom, mutational spectrum and mean target read depth. D, similarity matrix showing the cellular prevalence of all SNVs called in the tumor bulk (union of all reads from all regions). The cellular prevalence was inferred using PyClone, and a similarity matrix was created showing the clustering of the different SNVs with similar prevalence in 9,000 iterations.
The mutational landscape is the result of multiple mutational processes caused by different endogenous or exogenous damaging agents, as well as defects in the DNA repair or replicative mechanisms (37). The mutational signatures resulting from these processes have been characterized mathematically in human cancers (6, 35) and associated to known chemical, physical, or endogenous agents, such as UV, smoking, and ongoing deamination. Especially, the APOBEC mutational signature has been studied intensively (29), and APOBEC-related mutations have been observed at an increasing rate over time in lung cancer (38). Furthermore, the expression of APOBEC proteins has previously been associated with poor prognosis in breast cancer (39) and bladder cancer (8). The APOBEC mutational signature was recently shown to be significantly associated with high-risk NMIBC (11), and APOBEC mutations may be responsible for driving the tumor diversification and adaptation under selective pressure (10). We observed that APOBEC mutational signatures were more frequent in PD patients; however, heterogeneity in signature frequency during disease development was observed, indicating late tumor-specific (nonfield disease) APOBEC mutational activity.

Exome sequencing has identified genomic alterations in the Notch and Hedgehog signaling pathways to be important for tumor progression (40, 41). Furthermore, TP53 and FGFR3 mutations are present in tumors of different stages and partly mutually exclusive pathways (42). We found no significant difference in mutational load when comparing early-stage tumors from patients with different outcome. This may be caused by the relatively few samples analyzed, but it may also indicate a significant heterogeneity at the genomic level, and several pathways of progression may exist.

Cancer tissue generally shows transcriptional bias compared with normal tissue, due to a higher mutational load, leading some alleles to become transcriptionally silent. Five genes (PLEC, ACTG2, SVIL, COL5A1, and MMP2) showed significant ASE. These genes encode proteins associated with cell structure and motility. Plakins (PLEC) and actins (ACTG2) are important for cell motility and cytoskeleton integrity. Sup ervillin (SVIL) is tightly associated with actin filaments, and COL5A1 encodes a collagen subunit present in extracellular matrix (43). Finally, matrix metalloproteinases (MMP2) have the ability to cleave components of the extracellular matrix and are important for

---

**Figure 6.** Temporal and clonal heterogeneity in two bladder cancer samples. A, sample case: patient P26. Trinucleotide mutational context for each sample from P26. Each bar represents the relative ratio of a given mutation in a given trinucleotide context (dark color, Cat 1 SNVs; light colors, all SNVs). B, the phylogenetic tree is shown (top left) with an arrow indicating where a new mutational process occurred in the tumor development. The time span between tumors is shown (top right), and an interpretation in terms of main subclones is given (middle right). Cellular prevalence of individual SNVs was inferred with PyClone and plotted for each comparison (bottom): first tumor versus second tumor (left plot), first tumor versus third tumor (middle plot), and second tumor versus third tumor (right plot). C and D, as A and B for sample case: patient P19.

---

Lamy et al.
metastasis (44). Nonsense and splice site mutations showed increased wild-type allele expression, while missense and silent mutations showed increased expression of the alternate allele. The increased ASE for nonsense mutations is in agreement with lowered steady-state level of cytoplasmic mRNA by nonsense-mediated decay (45). Furthermore, the higher ASE of Ta tumors from PD patients compared with NPD patients indicates biological differences even at this early stage.

Efficient targeted therapy should preferably target alterations in all tumor cells. Thus, it is important to clearly identify the clonal mutations. However, observing clonality within a single tumor may not necessarily imply clonality across synchronous or metachronous tumors. Here, we showed that analysis of paired tumors identified potential therapeutic targets that were of clonal origin, and consequently, targeted therapy may be able to eradicate the field disease and multiple tumors.

Disclosure of Potential Conflicts of Interest
No potential conflicts of interest were disclosed.

Authors’ Contributions
Conception and design: P. Lamy, I. Nordentoft, T.F. Orntoft, L. Dynskjæt
Development of methodology: P. Lamy, I. Nordentoft, P. Villesen, J.S. Pedersen

Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): J. Nordenfot, K. Birkenkamp-Demtroeder, M.B.H. Thomsen, J. Hedegaard, M. Borre, J.B. Jensen, S. Høyer, T.F. Orntoft, L. Dynskjæt

Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): P. Lamy, I. Nordenfot, M.B.H. Thomsen, P. Villesen, S. Vang, J.S. Pedersen, T.F. Orntoft, L. Dynskjæt

Writing, review, and/or revision of the manuscript: P. Lamy, I. Nordenfot, K. Birkenkamp-Demtroeder, M.B.H. Thomsen, P. Villesen, S. Vang, J. Hedegaard, M. Borre, J.B. Jensen, S. Høyer, J.S. Pedersen, T.F. Orntoft, L. Dynskjæt

Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): I. Nordenfot, K. Birkenkamp-Demtroeder, M. Borre, S. Høyer, T.F. Orntoft

Study supervision: I. Nordenfot, J.S. Pedersen, T.F. Orntoft, L. Dynskjæt

Grant Support
This work was supported by The John and Birthe Meyer Foundation, The Lundbeck Foundation, The Danish Cancer Society, The Danish Cancer Biobank, and The Danish Research Council.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received February 17, 2016; revised May 31, 2016; accepted June 26, 2016; published OnlineFirst August 3, 2016.

References


Paired Exome Analysis Reveals Clonal Evolution and Potential Therapeutic Targets in Urothelial Carcinoma

Philippe Lamy, Iver Nordentoft, Karin Birkenkamp-Demtröder, et al.


Updated version
Access the most recent version of this article at: doi:10.1158/0008-5472.CAN-16-0436

Supplementary Material
Access the most recent supplemental material at: http://cancerres.aacrjournals.org/content/suppl/2016/08/03/0008-5472.CAN-16-0436.DC1

Cited articles
This article cites 44 articles, 8 of which you can access for free at: http://cancerres.aacrjournals.org/content/76/19/5894.full#ref-list-1

Citing articles
This article has been cited by 4 HighWire-hosted articles. Access the articles at: http://cancerres.aacrjournals.org/content/76/19/5894.full#related-urls

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

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

Permissions
To request permission to re-use all or part of this article, use this link http://cancerres.aacrjournals.org/content/76/19/5894. Click on "Request Permissions" which will take you to the Copyright Clearance Center's (CCC) Rightslink site.