Circulating tumor DNA reveals clinically-actionable somatic genome of metastatic bladder cancer

Gillian Vandekerkhove1*, Tilman Todenhöfer1,2*, Matti Annala1,3*, Werner J. Struss1, Amanda Wong1, Kevin Beja1, Elie Ritch1, Sonal Brahmbhatt1, Stanislav V. Volik1, Jörg Hennenlotter2, Matti Nykter3, Kim N. Chi4, Scott North5, Arnulf Stenzl2, Colin C. Collins1, Bernhard J. Eigl4, Peter C. Black1, Alexander W. Wyatt1.

1Vancouver Prostate Centre, Department of Urologic Sciences, University of British Columbia, British Columbia, Canada; 2Department of Urology, University of Tübingen, Tübingen, Germany; 3Institute of Biosciences and Medical Technology, University of Tampere, Tampere, Finland; 4Department of Medical Oncology, British Columbia Cancer Agency; 5Cross Cancer Institute, Department of Oncology, University of Alberta, Alberta, Canada; *indicates co-first authorship.

Running title: ctDNA alterations in metastatic bladder cancer

Keywords: ctDNA, cell-free DNA, cfDNA, liquid biopsy, muscle-invasive bladder cancer, precision oncology

This study was primarily supported by Bladder Cancer Canada (Alexander W. Wyatt, Bernhard J. Eigl, Peter C. Black). Additional support was provided by the Canadian Urologic Oncology Group (Peter C. Black), the Emil Aaltonen Foundation (Matti Annala), the Academy of Finland #269474 (Matti Nykter) and the German Cancer Foundation (Deutsche Krebshilfe; Tilman Todenhöfer).

Correspondence to Dr. Alexander Wyatt or Dr. Peter Black, Vancouver Prostate Centre, Department of Urologic Sciences, University of British Columbia, 2660 Oak Street, Vancouver, BC, V6H 3Z6, Canada; Tel: +1-604-875-4818; Fax: +1-604-875-5654; Email: awyatt@prostatecentre.com or peter.black@ubc.ca.

No relevant conflicts of interest to declare.
Abstract

Purpose: Targeted agents and immunotherapies promise to transform the treatment of metastatic bladder cancer (BCa), but therapy selection will depend on practical tumor molecular stratification. Circulating tumor DNA (ctDNA) is established in several solid malignancies as a minimally-invasive tool to profile the tumor genome in real-time, but is critically under-explored in BCa.

Experimental Design: We applied a combination of whole exome sequencing and targeted sequencing across 50 BCa driver genes to plasma cell-free DNA (cfDNA) from 51 patients with aggressive BCa, including 37 with metastatic disease.

Results: The majority of metastatic patients, but only 14% of patients with localized disease, had ctDNA proportions above 2% of total cfDNA (median 16.5%, range 3.9 - 72.6%). 12% of estimable samples had evidence of genome hypermutation. We reveal an aggressive mutational landscape in metastatic BCa with 95% of patients harboring deleterious alterations to TP53, RB1 or MDM2, and 70% harbouring a mutation or disrupting rearrangement affecting chromatin modifiers such as ARID1A. Targetable alterations in MAPK/ERK or PI3K/AKT/mTOR pathways were robustly detected, including amplification of ERBB2 (20% of patients) and activating hotspot mutations in PIK3CA (20%), with the latter mutually exclusive to truncating mutations in TSC1. A novel FGFR3 gene fusion was identified in consecutive samples from one patient.

Conclusions: Our study demonstrates that ctDNA provides a practical and cost-effective snapshot of driver gene status in metastatic BCa. The identification of a wide spectrum of clinically-informative somatic alterations nominates ctDNA as a tool to dissect disease pathogenesis and guide therapy selection in metastatic BCa.
Statement of translational relevance

Bladder cancer (BCa) is comprised of multiple molecular subtypes and a highly heterogeneous landscape of somatic mutations, copy number changes and chromosomal rearrangements. There is mounting evidence that tumor subtypes and distinct somatic alterations can influence patient response to existing and novel therapies for advanced BCa. As the arsenal of targeted and immunotherapies continues to grow, it is imperative that biomarker development hold pace to guide their implementation. Our results show that analysis of circulating tumor DNA (ctDNA) from the plasma of patients with metastatic BCa captures the somatic mutational profile in most patients. This suggests that ctDNA profiling can form a practical tool for real-time patient stratification, and thereby avoid the need for tissue biopsy or interrogation of potentially less-relevant archival tissue. Therefore, we propose that ctDNA collections be incorporated in future clinical studies of metastatic BCa, to better dissect therapy response and accelerate rational implementation of precision oncology.
Introduction

Urinary bladder cancer (BCa) is the fifth most common malignancy in North America, and was diagnosed in approximately 430,000 patients worldwide in 2012 (1). A quarter of patients present with muscle-invasive BCa, which is generally high-grade and progresses to metastatic disease in 50% of patients (2). Until recently, cisplatin-based chemotherapy was the only treatment with a proven survival benefit for metastatic BCa, and the 5-year survival rate with this regimen is only 5-15% (3). Recent data suggest that immune checkpoint blockade elicits responses in 15-30% of patients with advanced BCa that have progressed on cisplatin-based therapy (4–7). Furthermore, targeted agents directed against ERBB2 (HER2), FGFR3 or the PI3K/mTOR pathway may hold promise in patient subsets (8–10).

Patient molecular stratification will be critical for optimal use of second-line therapies in the metastatic setting. Although few metastatic tumors have been studied, primary BCa is molecularly heterogeneous and characterized by high somatic mutation rates (the third highest of all studied cancers) and recurrent copy number alterations. Common tumor suppressors (e.g. TP53, RB1, CDKN2A) are frequently lost and the majority of patients have alterations to chromatin modifiers and/or the PI3K/mTOR and MAPK pathways (11–15). Furthermore, there are distinct molecular subtypes based primarily on RNA expression (13,16–18). The influence of specific mutations or molecular subtypes on clinical outcomes is not yet clear, but there is evidence to suggest that ERCC2 and TSC1 mutations correlate with cisplatin and everolimus sensitivity respectively, and that RNA subtypes are linked to response to atezolizumab and cisplatin (4,16,19,20). High somatic mutation rates also appear to be a biomarker for response to immune checkpoint inhibition, consistent with similar data from melanoma, colorectal and non-small cell lung cancer (21–23). The majority of current biomarker approaches rely upon archival tissue obtained at transurethral resection or cystectomy, which may have reduced applicability in the setting of second and subsequent lines of therapy. A method for real-time genetic analysis of metastatic tumors is urgently needed.

The analysis of circulating tumor DNA (ctDNA) in plasma has shed light on the somatic landscape of metastatic disease in several solid malignancies. Prognostic and predictive ctDNA-based biomarkers are beginning to emerge in prostate, colon and pancreatic cancer (24–26). In non-small cell lung cancer, a ‘companion diagnostic’ ctDNA test for the EGFR T790M mutation recently received FDA approval. However, the abundance and utility of ctDNA in advanced BCa remains largely unexplored, with prior studies limited by small cohort sizes and dependency on single gene sequencing and/or digital droplet PCR (ddPCR) assays to detect specific missense mutations (27–29). While ddPCR provides the sensitivity required to detect extremely low levels of ctDNA (e.g. for predicting disease relapse after curative therapy (29), only broad next-generation sequencing assays can deliver a comprehensive analysis of clinically-relevant alterations such as PI3K/mTOR pathway deregulation or somatic hypermutation. In this study we applied a custom sequencing approach to plasma cell-free DNA (cfDNA) collected from a cohort of patients with aggressive BCa. We reveal the landscape of somatic alterations detected across 50 clinically-relevant driver genes in metastatic BCa, demonstrating the promise of liquid biopsies as both a discovery tool and a biomarker for therapy selection.
Materials and methods

Patient cohort and sample processing

We accrued 51 consecutive muscle-invasive BCa patients to our liquid biopsy program between December 2014 and December 2016 (Supplementary Table S1). Approval for this study was granted by the University of British Columbia Ethics Board, or the local ethics review board for participating sites in Edmonton, Canada and Tübingen, Germany. The study was conducted in accordance with the Declaration of Helsinki, and written informed consent was obtained from all participants prior to collection. Methods for plasma cfDNA and buffy coat fraction germline DNA (gDNA) extraction from whole blood are listed in the Supplementary Methods (30). Control DNA for testing the targeted sequencing panel was extracted from BCa cell lines RT4v6 and RT112. Cell lines were obtained in 2008 from MD Anderson Cancer Center and genotyped using a Promega StemElite kit to PCR amplify eight short tandem repeat loci plus a gender determining marker, Amelogenin. Mycoplasma testing was performed in July 2016 prior to DNA extraction using PlasmoTest (InvivoGen, San Diego, CA, USA).

Targeted and exome sequencing

We employed a targeted sequencing strategy capturing the exonic regions of 50 BCa relevant genes, as well as the TERT promoter and FGFR3 introns (Supplementary Table S2). Genes were selected based on the established landscape of recurrent mutations and copy number alterations in BCa (12–15). In total, 0.7 Mb of genomic sequence was captured (including off- and near-target regions). For each cfDNA sample, 10-25 ng of DNA was used for library preparation. A-tailing, end repair, Illumina-compatible adapter ligation and PCR amplification (between 15 and 17 cycles) was performed. Library quantification was carried out with the NanoDrop spectrophotometer, and each library was run on an ethidium bromide gel to confirm success. Up to 25 purified sample libraries at a time were multiplexed to obtain single pools with a combined mass of 1 ug, allowing a minimum 40 ng input for each sample library. These pools were hybridized to the capture panel for a minimum of 16 hours at 47°C. The subsequent wash, recovery and amplification of the captured regions was performed according to the NimbleGen SeqCap EZ system protocols. Final libraries were purified with Agencourt AMPure beads and quantitated using either the KAPA qPCR kit, or the Qubit 2.0 Fluorometer (Life Technologies) and Qubit dsDNA HS Assay Kit. Pools were diluted to 20 pM, and were sequenced on Illumina MiSeq (V3 600 cycle kit) or HiSeq 2500 (V4 250 cycle kit) machines. Preparation and sequencing of gDNA libraries was performed as described for cfDNA, with the following exceptions: all gDNA samples were sheared into 180 bp fragments with a Covaris focused-ultrasonicator prior to library preparation; 100 ng of DNA were used per library; PCR amplification was limited to 12 cycles. Exome capture and sequencing was performed using libraries previously prepared for targeted sequencing, and following the identical protocols described above, utilizing the Roche Nimblegen SeqCap EZ MedExome kit.
Alignment and quality control
Paired-end reads were aligned against the hg38 reference genome using Bowtie-2.2.4 (31). Optical and PCR duplicates were removed using samblaster-0.1.20 (32). Adapters were trimmed in paired mode using cutadapt-1.11 (33). Low-quality read tails (smoothed baseq < 30) were trimmed using an in-house algorithm. Per-base read coverages in target regions were counted using bedtools-2.25.0 (34). To verify sample identities and pairings, sample genotypes were compared at over 500 high confidence genotyped SNPs.

Somatic variant calling
Somatic mutations were called in cfDNA samples by searching for variants with a mutant allele fraction of ≥ 1%, and ≥ 10 supporting reads. Additionally, the allele fraction was required to be 25 times higher than the background error rate (i.e. the average allele fraction across all gDNA samples), and 3 times higher than the allele fraction in the paired gDNA sample. Protein-level consequences of variants were predicted using ANNOVAR (35), and all variants were annotated with their frequency in the COSMIC database. All variants were further annotated with their population frequency in the ExAC and KAVIAR databases, and somatic mutation candidates with a population frequency > 0.5% were discarded if paired gDNA exhibited low coverage at the site. Finally, all somatic mutation candidates were inspected in Integrative Genomics Viewer (IGV). Variants in low mappability regions and variants near indels were flagged for detailed scrutiny. Due to low coverage at TERT promoter mutation hotspots, these mutations were called by manual inspection of read pileups. PCR and Sanger sequencing were performed to confirm high allelic frequency somatic mutations in cases where flanking regions were amenable to primer design, and ample cfDNA was available.

The somatic mutation rate of a given sample (in mutations per megabase) was estimated as
\[ \sum_k \frac{10^6}{B_k}, \]
where \( B_k \) was the number of genomic locations sequenced to sufficient depth for mutation \( k \) to be detected. The threshold for sufficient depth was \( R / F_k \), where \( R \) was the minimum number of alternate allele reads required for a mutation to be called, and \( F_k \) was the allele fraction of mutation \( k \).

Estimation of ctDNA fraction
ctDNA fractions were estimated based on the allele fractions of autosomal somatic mutations. We first observed that the mutant allele fraction in diploid chromosomes is highest when the mutation is combined with loss-of-heterozygosity (somatic mutations in genes with a detectable amplification were omitted from ctDNA fraction estimation). In this situation, the mutant allele fraction (MAF) and ctDNA fraction are related as
\[ \text{MAF} = \frac{(\text{ctDNA} * 1)}{[(1 - \text{ctDNA}) * 2 + \text{ctDNA} * 1]}, \]
and so \( \text{ctDNA} = \frac{2}{(1 / \text{MAF} + 1)}. \) Since we could not detect deletions or loss-of-heterozygosity from low ctDNA fractions, we conservatively assumed that all mutations could have associated LOH. To deal with stochastic variation in observed mutant allele read counts, we modeled the mutant read count as arising from a binomial distribution, and conservatively calculated what the true mutant allele fraction would be if the highest observed mutant allele fraction was a 95% quantile outlier.
Copy number calling

Reads were counted in all capture regions using bedtools-2.25.0 (34). Coverage logratios were calculated against a low-noise median reference derived from all gDNA samples. GC fraction was calculated for all target regions, and Loess regression was applied to bait logratios to correct for GC content bias. After GC bias correction, read counts were corrected for residual differences in overall coverage using median-of-ratios normalization.

Since hemizygous copy number changes are associated with a loss of one allele at heterozygous bases, we used gDNA samples to find all heterozygous germline SNPs in captured regions, and then calculated the allele fractions of those SNPs in the corresponding cfDNA samples. We then calculated a median divergence from heterozygosity, defined as median(abs(AF - 0.5)), for each gene and cfDNA sample.

A deletion was called for a gene when coverage logratio ≤ -0.1 and MAF ≥ 0.6, or regardless of MAF if coverage logratio ≤ -0.3. A gain was called for a gene when coverage logratio ≥ 0.1 and MAF ≥ 0.6, or regardless of MAF if logratio ≥ 0.3. These conservative thresholds were determined empirically by studying a plot of coverage logratios and heterozygous SNP allele fractions in samples with and without detectable ctDNA (Supplementary Figure S1). Amplifications and homozygous deletions were called using very conservative logratio thresholds of 0.7 and -1.0, respectively.

cfDNA fraction corrected absolute gene copy numbers were estimated using the formula $C_{abs} = (2 * 2L - 2) / F + 2$, where $C_{abs}$ is the absolute copy number, $L$ is the median coverage logratio of targeted regions inside the gene, and $F$ is the cfDNA fraction.

Analysis of chromosomal rearrangements

Although most large deleterious genomic rearrangements fall within introns and are likely undetectable by an exon-limited panel, we searched for those in or near exonic regions by splitting unaligned reads into two 30 bp anchors that were then aligned to the hg38 genome. Discordant anchor pairs were grouped into clusters by position, and anchor pair clusters with 8 or more supporting reads were flagged as rearrangement candidates and manually curated using IGV and BLAT. For the FGFR3-ADD1 fusion gene, putative coiled coil oligomerization domains in the ADD1 protein were identified using the COILS software (36).

Results

High ctDNA burden in metastatic bladder cancer

To span the heterogeneous clinical spectrum of aggressive BCa we collected 85 plasma samples from 51 patients, including 14 patients with localized muscle-invasive BCa (at initial sample collection) and 37 with nodal and/or distant metastatic disease (Figure 1). Two patients provided matched metastatic tissue samples (peritoneal carcinomatosis and lymph node). Our cohort included plasma collected from patients prior to, during, and after platinum-based
chemotherapy in both the neoadjuvant/adjuvant localized and metastatic setting. We obtained significantly less cfDNA from chemotherapy-naive localized patient samples (n=13) compared to those with metastatic disease not currently receiving systemic therapy (n=42) (median 6.4 vs 20.3 ng/mL plasma; mean 9.6 vs 43.2; p=0.00038, rank-sum test) (Figure 2A). Samples from patients receiving platinum-based chemotherapy at time of collection (n=26) had high cfDNA yields (median 20.8 ng/mL plasma, mean 30.9, range 5.5 - 110) regardless of disease state (Figure 2A).

For cost-effective next-generation sequencing of cfDNA it is not possible to retain high sensitivity for ultra-low ctDNA fractions while simultaneously covering a large genomic territory. However, while the proportion of cfDNA that is tumor-derived (the ctDNA fraction) is largely unknown in metastatic BCa, it is often greater than 1% in advanced solid malignancies studied to date (24,37,38). In data from The Cancer Genome Atlas (TCGA), 98% of primary muscle-invasive BCa had a non-synonymous somatic mutation in at least one of 50 BCa genes in our panel (13). As such, we reasoned that with a sequencing depth of 500-1000x, somatic mutations would theoretically be detected in almost all patients should their ctDNA fraction exceed 5%. To test this, we spiked genomic DNA from BCa cell line RT4v6 into gDNA from patient 10044. Read support for five tested variants was observed at the lowest level of dilution (1% cell line DNA) (Supplementary Figure S2). Using our conservative thresholds for somatic variant calling, all five variants were independently detected at 10% dilution, 4 of 5 at 5% dilution, and 2 of 5 at 2% dilution. Strong copy number alterations such as CDKN2A biallelic loss and PIK3CA amplification (both present in RT4v6 cells) passed our logratio thresholds for detection at dilution factors of 20 - 30%, but not at lower dilutions, although we could not use germline SNPs to augment copy number calling as we would for paired patient specimens (see Methods) (Supplementary Figure S3).

We applied our targeted sequencing methodology to 62/85 cfDNA samples (median depth 753x, IQR 640x - 922x; Supplementary Figures S4 and S5), and matched leukocyte gDNAs for all patients subjected to cfDNA sequencing (median depth 427x, IQR 357x - 496x). This included all cfDNA samples (n=52/52) from metastatic patients together with representative samples (n=10/33) from localized muscle-invasive BCa patients. As expected, given their lower tumor burden, only 1/7 localized patients (3/10 samples, all originating from the same patient) had evidence of ctDNA above our detection thresholds (Figure 2B-D). Among patients diagnosed with metastatic disease, 58% of cfDNA samples (30/52; 24/37 patients) had evidence of ctDNA above 2% of total cfDNA, however 12 of these samples were collected from patients currently undergoing chemotherapy and/or without evidence of disease at time of collection. When considering samples obtained from patients with distant metastatic disease not currently receiving systemic therapy, 73% (16/22) had ctDNA fractions above our detection threshold (Figure 2C-D).

Among the 33 samples with evidence of ctDNA, the ctDNA fraction (see Methods) ranged from 3.9% to 72.6% (mean 26.6%, median 19.8%) (Figure 2B, Supplementary Table S3). Eleven samples had a ctDNA fraction over 30%, and all were from patients with distant metastatic disease at the time of collection. Conversely, 12 of 23 samples with ctDNA fractions between
3.9% and 30% were from patients with distant metastatic disease (p = 0.0058 for comparison to 11/11; Fisher’s exact test). In patients with distant metastatic disease, regardless of current treatment modality, there was no apparent difference in ctDNA fraction between samples from patients who had received prior cystectomy (n=21), and those where the primary tumor remained intact (n=12) (median 13.1% vs 6.8%, p=0.29, rank-sum test) (Figure 2D). Although these numbers are small, it suggests that the majority of ctDNA is being shed from metastatic sites in these patients.

Four cfDNA samples carried an estimated ctDNA fraction higher than 70% (Figure 2B). In each of these samples, the high ctDNA fraction was supported by multiple high allele fraction somatic mutations (Figure 2B) and strong copy number logratios (Supplementary Figure S6). All four patients also had distant metastases, and high cfDNA yields (between 60 - 369 ng/mL plasma). Sanger sequencing confirmed the presence of high allele fraction somatic mutations (Supplementary Figure S7).

Clinically-relevant alterations are detected in all patients with ctDNA

Across the 26 patients with quantifiable ctDNA in at least one cfDNA sample (or in two cases, metastatic tissue), we detected 281 somatic mutations including 121 protein altering mutations (Figure 3A, Supplementary Table S4). We identified 22 missense or truncating mutations in TP53 in 17/26 patients (65%; Figure 3B). Incorporating copy number results, 24/26 patients carried either a TP53 inactivating change (n=19), RB1 inactivating change (n=8), MDM2 gain (n=2), or CDKN2A loss (n=6), likely resulting in disrupted cell cycle regulation (Figure 3C). For the two patients without alterations in these four genes, E-006 had non-metastatic disease, and BC-008 had a low ctDNA fraction not amenable to copy number analysis. Four patients exhibited definitive evidence for biallelic inactivation of RB1 or CDKN2A (Figure 3C). Over half the cohort (19/26 patients) had mutations or disrupting rearrangements in chromatin modifier genes (Figure 3C), including eight truncating mutations within ARID1A that remove the DNA binding domain and/or glucocorticoid receptor binding domain, and eight truncating mutations in KMT2D (MLL2). TERT promoter mutations were identified in 12/26 patients: ten patients harbourved the chr5:1295113:G>A mutation (reported in 65% of BCa), while two carried the chr5:1295135:G>A mutation (reported in 10% of BCa) (39). The majority of patients had alterations to the PI3K/mTOR pathway, including six with hotspot missense mutations in PIK3CA (K111E, E542K, E545K (n=2), E674Q or E726K), four with truncating mutations in TSC1 (mutually exclusive with PIK3CA hotspot mutations), one with PIK3R1 stopgain, one with PTEN stopgain, and one with a TSC2 truncating rearrangement (Figure 3C). Six patients showed evidence for PTEN or PIK3R1 deletion.

Nine patients (35%) carried ERBB2 activating somatic changes, including amplification in five patients (Supplementary Figure S8), and hotspot mutations (S310F, L755S, I767M, V777L) in four patients. Four patients carried ERBB3 activating alterations, including two amplifications and two hotspot mutations (M91I, V104L) (Figure 3C). Remarkably, one patient had an average of 71 copies ERBB2 in his ctDNA. Three more patients carried activating RAS mutations (KRAS G12D (n=2), HRAS Q61R), and one patient carried a KRAS amplification. In total, 15/26 patients carried activating somatic alterations in the MAPK pathway.
Overall we observed no significant differences in mutation frequency (across the 50 genes in our panel) between samples in our metastatic cohort and in the TCGA localized muscle-invasive BCa cohort (Supplementary Figure S9) (13). Even the distribution (within each gene) and nature of mutations (e.g. predominantly missense versus truncating) were highly consistent between the two cohorts, providing a strong measure of validation for mutations described here.

Two patients (T-015 and T-019) had both metastatic tissue and matched plasma ctDNA sequenced, but only T-015 had quantifiable ctDNA (estimated 5% fraction). Reassuringly, the solid and liquid biopsies for T-015 revealed the same five somatic mutations, although the CDKN2A missense mutation was detected at a variant allele frequency of only 0.9% in the liquid biopsy (Supplementary Figure S10). Four patients in our cohort had evidence of ctDNA in more than one plasma collection obtained over the course of chemotherapy. Temporal samples from the three patients with distant metastatic disease were highly consistent, with the same patient-unique mutations detected over multiple timepoints (Supplementary Figure S11).

Evidence of somatic hypermutation

Projecting our targeted sequencing data across the genome suggested a mean and median somatic mutation rate of 17.2 and 14.3 per Mb respectively (Figure 3A, Supplementary Table S4). 3/26 (12%) of patients exhibited projected rates ≥ 30 per Mb: close to that detected in mismatch repair deficient colorectal cancers (21). Our findings are higher than the mean and median somatic mutation rates of 7.7 and 5.5 per Mb respectively observed in primary muscle-invasive BCa (13).

To assess the degree to which our targeted sequencing derived mutation rates recapitulate those observed over the entire exome, we subjected 11 cfDNA libraries from 8 patients (the majority of cfDNAs harbouring a ctDNA fraction over 25%, and their matched gDNA libraries) to exome sequencing (median depth 176x, IQR 171x - 187x). Somatic mutation allele fractions and copy number coverage logratios observed in exome sequencing data were consistent with the ctDNA fractions derived from targeted sequencing data (Figure 4). When considering mutations robustly detectable by both panels (i.e. those with allele fraction > 10%) we observed a high correlation for mutation counts between the 50-gene targeted panel and whole exome sequencing data (Pearson correlation = 0.83) (Supplementary Figure S12). Estimated somatic mutation rates (see Methods) were also correlated between the two panels (Pearson correlation = 0.71) (Supplementary Figure S12). Nevertheless, the targeted panel consistently overestimated somatic mutation rates relative to whole exome sequencing (median 66% increase in mutation rate). This increase may be due to recurrent driver mutations in the targeted genes, although we observed a high proportion of synonymous or non-coding variants (unlikely to be under selective pressure) in the samples with the highest mutation rates (Supplementary Figure S12B).

At the individual sample level, the exome data provided strong technical validation for mutation calls from the initial targeted sequencing data (Supplementary Figure S12). The higher depth of sequencing in the targeted panel data inevitably detected more subclonal mutations than...
exome sequencing. For example, BC-005 (predicted ctDNA fraction = 26%) had 25 somatic mutations detected in targeted sequencing data (between AF 2.0 - 17.3%). Only four of these mutations were detected in whole exome sequencing with an allele fraction of 10% or higher (the detection threshold in exome data) (Supplementary Figure S12C). None of the 25 mutations showed a significantly different allele fraction between targeted and whole exome sequencing (p > 0.01 for all, Fisher's exact test). This suggests that the subclonal (< 10% allele fraction) mutations detected by targeted sequencing are real, but only detectable at high sequencing depth, highlighting the impact of sequencing depth on estimated somatic mutation rates.

Although the heterogeneity of our cohort prevents broad analysis of clinical outcomes, we note that six metastatic patients in our cohort had received immunotherapy during the course of their disease (T-017, T-014, T-009: nivolumab; BC-017: atezolizumab; BC-014: durvalumab; 10044: pembrolizumab). BC-017 experienced a complete response to atezolizumab two years prior to plasma sampling (treatment was ceased due to immune-related colitis). T-014 had a partial response to nivolumab. Both patients exhibited high projected somatic mutation rates (22 and 39 Mb⁻¹, respectively). The remainder (all non-responders) had either no/low ctDNA (T-017, T-009) or a low/median mutation rate (BC-014, 10044).

Identification of novel fusion genes and complex gene rearrangements

We did not detect any FGFR3 mutations, consistent with their predominance in non-muscle invasive papillary BCa (13). However, we identified a novel FGFR3-ADD1 fusion event in two cfDNA samples (collected 8 months apart) from patient T-002 (Figure 5; Supplementary Table S5). The rearrangement juxtaposes the 3'-UTR of FGFR3 with the second intron of ADD1, likely resulting in an in-frame chimeric protein where the last FGFR3 exon is spliced out due to its lack of a splice donor site. Since ADD1 contains a dimerization domain, it plausibly enables constitutive activation of FGFR3, functioning similarly to the well-described FGFR3-TACC3 fusion gene (40). We did not detect FGFR3-TACC3 in any patients here, although it was robustly identified in two cell lines (RT112 and RT4v6; Supplementary Figure S13). In addition, we report two different ERBB2 fusions in patients with ERBB2 gene amplification. The fusion genes, ERBB2-CDK12 and MED1-ERBB2, likely represent secondary passenger events accompanying the formation of high-level gene amplification. Also detected (in four separate patients) were a 27 kb deletion in ARID1A, 34 bp deletion overlapping a TSC2 splice donor site, and truncating rearrangements in RB1 and RUNX3 (Supplementary Table S5). Repeat samples from the same patient, when available, supported the rearrangements.

Discussion

Practical molecular stratification of patient tumors will be critical to guide the use of emerging targeted therapies and immunotherapy in metastatic BCa. However, the limited availability of tissue from metastases in routine clinical practice hampers attempts to implement real-time precision oncology in this setting. In this context, we describe clinically-relevant somatic alterations detected through minimally-invasive plasma ctDNA profiling in the majority of patients with metastatic BCa.
Confidence in our results can be drawn from several sources of evidence. Firstly, the distribution and frequency of mutations, copy number alterations and chromosomal rearrangements was highly consistent with localized muscle-invasive BCa (13,14). Secondly, the only recurrent mutations detected in our study were known hotspots in \textit{PIK3CA}, \textit{ERBB2} and the \textit{TERT} promoter: all other mutations were unique to individual patients. Thirdly, when multiple temporal samples from individual patients were analyzed, the same mutations were consistently observed. Fourth, somatic mutations with high allelic frequency from our targeted sequencing approach were independently verified by Sanger sequencing and matched exome sequencing. Copy number calls and ctDNA fractions derived from matched exome sequencing data were also highly consistent with results from our targeted panel. Finally, similar to observations from other malignancies (28), we observed higher ctDNA fractions in patients with distant metastatic disease than those with local lymph node involvement or localized cancer.

Prior studies have suggested that BCa releases abundant ctDNA, but were based on very small patient numbers and assays limited to specific missense mutations or single genes (27–29). In our study, the majority of metastatic patients had ctDNA fractions above 2%, and a third had fractions greater than 20%. This implies that samples from most patients with progressive metastatic disease are suitable for cost-effective targeted sequencing approaches, and a significant proportion for standard exome sequencing. However, even when excluding clinically metastatic patients likely to have a very low or absent tumor burden (e.g. those with pathologic lymph node disease) there remained some plasma ctDNA samples from metastatic patients that yielded no ctDNA evidence. Our assay was also relatively uninformative in the localized setting, although the availability of tumor tissue collected at transurethral resection (or cystectomy) and urine-based biomarkers reduce the rationale for pursuing plasma ctDNA in this setting (27,41). Nevertheless, future studies seeking to capture somatic information from all metastatic patients could augment targeted panel sequencing with digital droplet PCR for patient-specific mutations identified through archival tissue analyses or for recurrent hotspot mutations (e.g. in \textit{PI3KCA} or \textit{TERT}), as has been applied successfully to clinically-localized BCa patients (27,29). Alternatively, the incorporation of unique molecular identifiers during library preparation and subsequent ultra-deep sequencing has proved successful in other solid malignancies for detecting mutant alleles as rare as 0.02% (28,42), but this can become expensive if applied beyond a few genes.

We observed high total cfDNA yields in patients receiving chemotherapy, in both the localized and metastatic setting. Since this was not accompanied by increased (or any) ctDNA in most instances it suggests elevated non-malignant cell death in the wake of systemic cytotoxic therapy. In pregnant women, fetal cfDNA is cleared rapidly from circulation after childbirth (43), and ctDNA levels reduce after surgery in patients with non-small cell lung cancer and colorectal cancer (44,45). The same is likely true in BCa patients responding to chemotherapy. Future ctDNA studies in BCa should carefully consider the timing of plasma collections, and be cautious extrapolating data from other cancers where overall cfDNA yield correlates with tumor burden (44).
In clinical trials to date, approximately 15-30% of BCa patients appear to respond to checkpoint blockade (2,4–7). In the atezolizumab phase II trial, responding patients had a significantly higher mutation rate than non-responders (4). In our study, approximately 12% of metastatic patients had projected mutation rates (from targeted sequencing) above 30 per Mb: levels of hypermutation associated with response to immunotherapy in other cancers (21–23). It is likely that our targeted approach, encompassing frequently mutated genes, overestimates global mutation rates. However, overall trends in mutation counts from targeted sequencing were consistent with those derived from matched exome sequencing, suggesting that it is at least possible to distinguish between relatively high and low mutational burdens. Certainly, a minimally-invasive assay for approximate mutational burden, as described herein, could prove valuable to the dissection of patient response to immune checkpoint blockade (4). It is plausible that ctDNA could be used in combination with immunohistochemistry (IHC) of PD-1/L1 and molecular subtype to help select patients most suitable for treatment.

The practical detection of mutations, copy number changes and even novel fusion genes from plasma in metastatic BCa offers exciting possibilities for precision oncology. We identified 11 patients with amplification and/or activating hotspot mutation to ERBB2/ERBB3. A recent phase II trial evaluated afatinib (an irreversible inhibitor of the ERBB family) in BCa and found that activity was confined to patients with alterations in ERBB genes (46). The future of other HER2-directed agents such as trastuzumab and lapatinib in BCa also likely depends upon biomarker-driven patient selection (10). Clinical trials of FGFR3 inhibitors are underway in BCa, but dovitinib (for example) has shown no activity to date in patients with FGFR3 mutations or positive immunohistochemistry (47). FGFR3 inhibitors are not yet tested in patients known to harbour activating gene fusions, even though BCa cell lines expressing FGFR3 fusions are extremely sensitive to FGFR-targeted agents, more so than those carrying mutations (40). Our data demonstrate that ctDNA profiling of FGFR3 introns alone is sufficient to identify novel FGFR3 fusions. We identified several patients with TSC1/TSC2 truncating alterations, which are reported to confer sensitivity to the mTOR inhibitor everolimus (19). Indeed, the PI3K/AKT/mTOR pathway was aberrant in a large proportion of patients, suggesting opportunities for testing of AKT inhibitors such as MK2206 that has previously demonstrated activity in BCa cell lines (14). We detected several mutations in DNA repair genes, including a missense mutation in ERCC2 and two truncating ATM mutations, raising the possibility that ctDNA profiling may also help guide implementation of cisplatin-based chemotherapy (48,49).

The profiling of plasma ctDNA also shows promise for monitoring therapy response or disease progression in BCa patients. Here, we demonstrate proof-of-principle for ctDNA profiling to capture temporal ctDNA fraction and the prevalence of driver mutations in three patients with distant BCa metastases undergoing chemotherapy. In the future this approach could be leveraged to dissect associations between changes in ctDNA fractions during treatment and clinical outcomes. Indeed, a recent study suggested that presence of ctDNA in plasma samples collected post-cystectomy in clinically-localized BCa was associated with disease recurrence (29), which is consistent with breast cancer where ctDNA detection after curative therapy portends rapid metastatic relapse (50). An added advantage of ctDNA profiling is the theoretical potential to survey intra-patient tumor heterogeneity, and the effect of the therapy on clonal...
dynamics. Advanced BCa was recently shown to be characterized by high levels of intra-patient heterogeneity (51), but the comprehensive dissection of longitudinal heterogeneity in ctDNA samples will likely require broader genome-wide approaches and more matched metastatic tissue, than used in the present study.

Our study demonstrates that the liquid biopsy can provide a practical and cost-effective snapshot of BCa driver genes status in patients with advanced disease. Given the relative ease of plasma collection, we propose that ctDNA analysis be incorporated in future clinical studies of metastatic BCa, where possible. With the proviso of comprehensive somatic data the liquid biopsy can inform valuable context to better dissect therapy response and accelerate rational implementation of precision oncology.
References


Figure legends

Figure 1. Overview of the bladder cancer patient cohort.

Figure 2. Association of cell-free DNA (cfDNA) yields and circulating tumor DNA (ctDNA) fractions with clinical characteristics. (A) Relationship of cfDNA yield to presence of metastases and concurrent systemic therapy (rx) at time of sample collection. (B) Estimated ctDNA fractions for all samples (gray bars, left y-axis), and supporting evidence. Each dot represents a somatic mutation. Vertical position of dots indicates mutant allele fraction (right y-axis). (C) Relationship of ctDNA fraction to presence of metastases and ongoing systemic therapy at time of sample collection. (D) cfDNA yields and ctDNA fractions across the cohort, grouped by disease state and chemotherapy exposure at the time of sample collection.

Figure 3. Landscape of somatic alterations detected in circulating tumor DNA (ctDNA) from patients with advanced bladder cancer. (A) Integrated mutation, copy number, and chromosomal rearrangement results from ctDNA positive patients. Where multiple collections from the same patient had evidence for ctDNA, the sample with the highest ctDNA fraction is shown. The top panel provides the ctDNA fraction per plasma sample (or percent tumor-derived DNA for the two tissue samples). Results from all 50 genes in the targeted panel are shown, sorted by chromosomal position. The bar plot on the right demonstrates the frequency of somatic alteration for each gene in the targeted panel, among the 25 patients with ctDNA. (B) Lollipop plots for three highly mutated genes, demonstrating the location of mutations relative to protein domains. (C) Oncoprints showing recurrent disruption of key pathways associated with bladder cancer development and progression.

Figure 4. Concordance of circulating tumor DNA (ctDNA) results between targeted and whole exome sequencing. (A-B) Mutant allele fractions from matched targeted (A) and exome sequencing (B) showing high inter-platform similarity and support for estimated ctDNA fractions. Each dot represents a single somatic mutation. In (A) the bar plot indicates predicted ctDNA fractions, while in (B) the grey clouds represent kernel density estimates of mutant allele fractions. (C-D) Exome sequencing derived copy number coverage logratios and germline SNP allele fractions for BC-002 (C) and T-001-3rd (D). Note that single copy deletions and gains (in the logratio plots) are robustly supported by deviations in heterozygous germline SNP allele fractions. The histograms on the far right indicate density of coverage logratios across the exome showing distinct copy number states. Red, white, and blue circles indicate where single copy gain, copy neutral state, and hemizygous deletion peaks are expected to fall given the ctDNA fractions shown above the vertical line. Note that the ctDNA fractions estimated from whole exome sequencing are very close to the ctDNA fractions estimated from targeted sequencing.

Figure 5. Novel FGFR3-ADD1 gene fusion identified in multiple cell-free DNA (cfDNA) samples from patient T-002. (A) Graphical representation of the fusion gene, predicted fusion pre-mRNA transcript, and chimeric in-frame FGFR3-ADD1 protein. (B) Unique cfDNA sequencing reads
supporting the fusion junction. A total of 55 unique supporting DNA fragments were found in the T-002-1st-cfDNA sample.
Figure 4

(A) Targeted 50-gene panel

Mutant allele fraction (%)

90 80 70 60 50 40 30 20 10 0

Somatic mutations
- TERT promoter
- TP53 protein altering
- ARID1A protein altering
- Other protein altering
- Other silent

(B) Whole exome sequencing

Mutant allele fraction (%)

10 20 30 40 50 60 70 80 90

- T-018-1st
- BC-002-1st
- T-001-3rd
- T-002-1st
- T-001-1st
- BC-003-1st
- T-001-2nd
- 10044-2nd
- T-002-3rd
- BC-005-1st
- BC-009-1st

(C) (D)
Figure 1

[Image of a bar chart showing samples across different disease states, sex, and age categories. The chart includes bar lengths labeled with sample identifiers and colors indicating different states, such as localized, pathological LN metastases, and clinical LN metastases. The sex is indicated with red for female and blue for male, and the age is divided into 45-60, 61-70, and 71-80 categories.]
Circulating tumor DNA reveals clinically-actionable somatic genome of metastatic bladder cancer

Gillian R Vandekerkhove, Tilman Todenhöfer, Matti Annala, et al.

Updated version
Access the most recent version of this article at:
doi:10.1158/1078-0432.CCR-17-1140

Author Manuscript
Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.

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, contact the AACR Publications Department at permissions@aacr.org.