

Detailed overview of patients disease courses and analyzed samples. Lines represent patient's disease courses from time of first analyzed tumor or urine sample to end of FU centered around the first induction treatment of BCG with at least five instillations. Colors indicate performed analyses. Shapes indicate the number of analyses per clinical timepoint. Left: full FU. Right: Two years pre and post-BCG illustrated.

Lower exhaustion Higher exhaustion Excessive exhaustion

Study and sample overview and RNAseq based analyses. A) Overview of study with disease course including BCG treatment of at least five instillations and sample types at different time points. Dashed lines indicate samples analyzed from a limited number of patients. Design of custom panel for deep targeted sequencing is outlined. Created with BioRender.com (publication license obtained). B) Venn diagram showing overlapping samples between analysis methods. C) Association between UROMOL2021 tumor subtype and time from BCG treatment for all post-BCG samples sequenced for the project (Wilcoxon Rank Sum test). D) CD8 adjusted exhaustion status in BCG-responsive and BCG-failure patients in pre- and post-BCG tumors, respectively (Fisher's Exact test). E) Correlation between CD8 T-cell score and exhaustion score. Green = lower exhaustion (residuals \leq =1.04). Purple = excessive exhaustion (residuals >1.04). Pearson correlation was used to determine the correlation coefficient *R* and *p*-value. F) Excessive exhaustion status in BCG-responsive and BCG-failure patients in pre- and post-BCG tumors, respectively (Fisher's Exact test).

Correlations between tumor WES characteristics and BCG-response. A) The number of single nucleotide variants (SNVs) in pre-BCG tumor samples correlated to response to BCG. B) The number of INDELs in pre-BCG tumor samples correlated to response to BCG. C) The number of APOBEC-related mutations (SBS2+SBS13) in pre-BCG tumor samples correlated to response to BCG. D) The number of SBS4-related mutations in pre-BCG tumor samples correlated to response to BCG. E) The number of SBS1-related mutations in pre-BCG tumor samples correlated to response to BCG. F) The number of SBS5-related mutations in pre-BCG tumor samples correlated to response to BCG. G) The number of SBS25-related mutations in pre-BCG tumor samples correlated to response to BCG. Green = BCGresponsive patients. Orange = BCG-failure patients. The Wilcoxon Rank Sum test was used for statistical comparisons.

Exhaustion status post−BCG Higher exhaustion Higher exhaustion

Correlations between post-BCG tumor exhaustion status and pre-BCG tumor WES characteristics. A) The number of single nucleotide variants (SNVs) in pre-BCG tumor samples correlated to post-BCG exhaustion status. B) The number of INDELs in pre-BCG tumor samples correlated to post-BCG exhaustion status. C) The number of APOBEC-related mutations (SBS2+SBS13) in pre-BCG tumor samples correlated to post-BCG exhaustion status. D) The number of SBS4-related mutations in pre-BCG tumor samples correlated to post-BCG exhaustion status. E) The number of SBS1-related mutations in pre-BCG tumor samples correlated to post-BCG exhaustion status. F) The number of SBS5-related mutations in pre-BCG tumor samples correlated to post-BCG exhaustion status. G) The number of SBS25-related mutations in pre-BCG tumor samples correlated to post-BCG exhaustion status. Pink = Higher post-BCG exhaustion. Turquoise = Lower post-BCG exhaustion. The Wilcoxon Rank Sum test was used for statistical comparisons.

Overview of identified mutational signatures. A) Detailed mutational profiles for the six mutational signatures identified by de novo extraction using the R package MutationalPatterns. Extracted signatures were named according to COSMIC signatures with the highest resemblance, as estimated using cosine similarity and resulting in the following values: $SBS1 = 0.91$, $SBS2$, 0.92 , $SBS4 = 0.89$, $SBS5 = 0.75$, $SBS13 = 0.76$, $SBS25 = 0.64$. B) Relative contribution of the identified mutational signatures. Sample clustering was performed based on euclidean distances.

Supplementary Table S1: Olink proteins and results

Overview of analyzed proteins using the Olink proteomics platform and associated biological processes. P-values and Benjamini-Hochberg adjusted P-values from paired ttests for comparisons between pre-BCG and post-BCG samples for all patients, BCGresponsive patients, and patients with BCG-failure are shown. Green color indicates pvalues < 0.05 and blue color indicates q-values <0.1.

Supplementary Table S2: Gene lists used for the RNA-based estimation of immune cell populations. Gene lists were obtained from¹⁴, except gene markers for CD4 T-cells, which were obtained from¹³.

Supplementary Materials and Methods

Clinical samples

Tumor biopsies were either dry-frozen or fresh frozen (FF) embedded in TissueTek OCTTM Compound (Sakura Finetek), snap-frozen in liquid nitrogen and stored at -80°C or obtained from formalin fixed paraffin embedded (FFPE) samples. Blood samples from all patients were stored in EDTA tubes at -80°C. Urine supernatants were centrifuged and stored at -80°C. Urine samples for protein analyses were sampled at a maximum of four months before and after BCG treatment.

DNA and RNA extraction

Tumor DNA from FF and dry-frozen tumors was extracted with Gentra Puregene Tissue Kit (Qiagen). DNA from FFPE samples was extracted using AllPrep DNA/RNA Kit (Qiagen). From peripheral blood, leukocyte DNA for germline (GL) reference was extracted using Qiasymphony DSP DNA midi kit (Qiagen) for all patients. Total Tumor RNA was extracted using RNeasy Mini and Micro Kits (Qiagen).

From tumors, haematoxylin and eosin (HE) stained sections of 4 μm were included to estimate carcinoma cell percentage.

DNA and RNA from tumors were extracted from serial cryosections of 20-25 μm. DNA from FFPE tumor specimens was extracted using punches taken from areas of the tumor with high estimated carcinoma cell percentage.

Urine (20–50 ml) was centrifuged at 3000G for 10 minutes, and supernatants were stored at –80C. Urine samples were thawed on ice, 10µl of urine was stored for Olink proteomics analyses (described below), PBS was added to remaining urine for a total volume of 4ml, and samples were centrifuged at 3000G for 10 minutes. Total cell-free DNA (cfDNA) was extracted from a median of 3.6 ml urine supernatants (range 0.7–3.65 ml). 10% ATL lysis buffer (median: 400µl, range: 100µl-500µl) (Qiagen, Aarhus, Denmark) was added, samples were incubated for 1.5-2h at room temperature followed by centrifugation at 1300G for 10 minutes. cfDNA extraction was performed using the QIAsymphony DSP Circulating DNA Kit (Qiagen). DNA was eluted in 60 µl ddH2O in DNA lobind tubes (Sigma-Aldrich).

Design of custom panel for targeted sequencing

For deep targeted sequencing of tumor specific mutations, we designed NGS panels where unique mutations (n=10-71) from 154 patients were included. Single nucleotide variants (SNVs) for the panels were selected from tumor guided WES analysis. Three panels covering mutations from 48, 54, and 52 patients, respectively, were designed.

SNVs were included based on the following criteria: 1. High or moderate impact mutations were favored. 2) Mutations with high variant allele frequencies (VAFs). 3) Mutations in known oncogenic genes from OnkoKB. 4) Bladder cancer associated genes (The Cancer Genome Atlas^{[1,2](https://paperpile.com/c/DEg4PA/ZAMZ+bdDl)}, MSKCC cBioPortal for Cancer Genomics^{[3,4](https://paperpile.com/c/DEg4PA/ZpaF+5cHl)}, Uromol 2021^{3,4}, and Lamy et al^{[5](https://paperpile.com/c/DEg4PA/9Jxd)}). 5) Mutations in the most exonically variable genes, reported by the Ingenuity Variant Analysis (IVA) software and/or at error-prone positions/in commonly reported erroneous contexts (C>T) were excluded, unless present in cancer driver genes or known bladder cancer genes. The three panels differed slightly in mutation selection.

Library preparation, deep-targeted sequencing and Mutation Calling of urinary

NGS libraries of cfDNA from urine supernatant were prepared using a modified version of the Mechanical Fragmentation Library preparation kit from Twist Bioscience in order to increase small fragment preservation and increase conversion rate. For robust error correction, 9 bp Unique Molecular Identifiers (UMI) were incorporated by replacing the twist adaptors with xGen™ UDI-UMI Adapters (Integrated DNA Technologies). The maximum cfDNA input was set to 100ng with no lower limit. Target enrichment was conducted using a modified version of the Twist Target Enrichment Protocol (Twist Bioscience) in combination with custom Twist panels (each SNV was captured by 120 bp probes on both strands). UMI consensus base calls were called using the fgbio tool package (v1.4.0). The ClipBAM (Hard-clipping) function of fgbio was enabled to avoid double count of variant calls of reads from the same template by clipping overlap between the reads. This was used for calculating UMI family size and UMI read depth. For UMI consensus collapsing, three reads were required as minimum and an editing distance of one was allowed. However, for secondary alignments, reads were not clipped. Realignment at INDEL positions were made using abra2 (v2.24). DeepSNV [6](https://paperpile.com/c/DEg4PA/QfHnm) was applied for detection and quantification of low-frequency tdDNA mutations in urine samples.

Calculations of UMI consensus read depth and UMI family sizes were based on the fgbio tool package (v1.4.0) that was run with the ClipBAM (Hard-clipping) function. Realignment at INDEL positions are made using abra2 (v2.24). For UMI consensus collapsing, three reads were required as minimum and an editing distance of one was allowed.

Reads were mapped against the hg38 genome using bwa mem v.0.7.17. Mapped reads were grouped and consensus reads were generated from UMIs. Hereafter, read counts for the positions included in the panel were evaluated using the pileup tool bam read count. Only mutations detected in both the 5'-3' direction and the 3'-5' direction were included. Furthermore, at least three reads with different UMIs supporting a given mutation were required.

Copies per ml were calculated as follows:

GE/ml = (DNA input(ng)*VAF*(1000 pg/ng)/(3.3pg/GE))/urine (ml)

Clearance of tumor derived DNA (tdDNA) was defined as VAF \neq 0 in pre-BCG samples and VAF=0 in post-BCG samples.

Neoantigen Calling

HLA types were called using POLYSOLVER^{[7](https://paperpile.com/c/DEg4PA/n1zp)}, xHLA^{[8](https://paperpile.com/c/DEg4PA/0uLm)}, and OptiType^{[9](https://paperpile.com/c/DEg4PA/Qizp)}, with patient HLA type decided by consensus vote supported by at least two algorithms. If no majority could be reached, POLYSOLVER was used. All novel 9-11mer peptide fragments were generated by MuPeXI^{[10](https://paperpile.com/c/DEg4PA/x5JU)} and eluted ligands (EL) rank-percentage scores for all HLA alleles were predicted by NetMHCpan-4.1^{[11](https://paperpile.com/c/DEg4PA/OYvL)}. The rank-percentage score represents the rank of the fragments EL probability compared to a set of random natural peptides. A mutation was considered a neoantigen if at least one fragment had an EL rank percentage score<2%, for at least one HLA allele.

DeepSNV

The inclusion of genomic positions associated with multiple patients on every custom panel, facilitates an abundance of sequencing data for every genomic position with no mutations expected to be present. We exploited this by employing an analysis framework based on a maximum likelihood implementation of the shearwater algorithm developed by Gerstung *et* al.^{[6](https://paperpile.com/c/DEg4PA/QfHnm)}. In brief, a background error model was built by fitting presumably non-mutated data, i.e. data from all samples not associated with a given mutation, to a binomial distribution with site-specific calculation of the dispersion for every mutation of interest. Samples with a VAF at a given position above 25% were excluded when generating the error model. Similarly, samples with a read depth below 50 at a given position were excluded when generating the error model. Furthermore, positions with an average error-rate above 10% across presumably non-mutated samples were excluded. Presumably mutated data, i.e. data from the target positions of a given sample, was assessed for a statistical significant difference compared to the background error model. Resulting *p*-values were corrected for multiple testing using the Benjamini-Hochberg procedure and adjusted *p*-values below 0.05 were considered significant.

In addition, to exploit that data were generated for all positions in a custom panel for every sample, we established an analysis to assess the combined signal from all target positions for a given sample. We used Fisher's method to calculate the sum of logs for all *p*-values obtained when analyzing the target positions. To gauge a randomly expected level of signal, we randomly sampled the same number of mutations as the target positions from non-target positions and similarly calculated the sum of logs. This was performed 10,000 times. Only samples with a sum of log scores greater than all scores from random sampling were considered ctDNA positive.

RNA sequencing data

Salmon^{[12](https://paperpile.com/c/DEg4PA/mPA8)} was used to quantify the expression of transcripts using annotation from the Gencode release 33 on genome assembly GRCh38.

Spiderplots were generated using the Radarchart function in the R package fmsb.

We estimated immune cell populations from the RNA-seq data using established gene expression signatures as in Rosenthal *et al. [13–15](https://paperpile.com/c/DEg4PA/AQJ6i+HYTTn+rkR3Y)* . A score for each immune cell population was calculated as the mean expression of all marker genes for the given cell type and a total immune score was defined as the mean of all immune cell population scores. Individual scores for B-cells, neutrophils and NK-cells were not evaluated as genes were missing for these cell population signatures. Gene lists are available in Supplementary Table S3.

WES analyses

MutationalPatterns was used for *de novo* extraction of mutational signatures [16,17](https://paperpile.com/c/DEg4PA/IAxc+LcGG)

References

- 1. Robertson, A. G. *et al.* Comprehensive Molecular [Characterization](http://paperpile.com/b/DEg4PA/ZAMZ) of Muscle-Invasive Bladder [Cancer.](http://paperpile.com/b/DEg4PA/ZAMZ) *Cell* **174**, 1033 (2018).
- 2. Cancer Genome Atlas Research Network. Comprehensive molecular [characterization](http://paperpile.com/b/DEg4PA/bdDl) of

urothelial bladder [carcinoma.](http://paperpile.com/b/DEg4PA/bdDl) *Nature* **507**, 315–322 (2014).

- 3. cBioPortal for Cancer [Genomics.](http://paperpile.com/b/DEg4PA/ZpaF) <https://www.cbioportal.org/>[.](http://paperpile.com/b/DEg4PA/ZpaF)
- 4. Lindskrog, S. V. *et al.* An integrated [multi-omics](http://paperpile.com/b/DEg4PA/5cHl) analysis identifies clinically relevant molecular subtypes of [non-muscle-invasive](http://paperpile.com/b/DEg4PA/5cHl) bladder cancer. *Urology* (2020) [doi:](http://paperpile.com/b/DEg4PA/5cHl)[10.1101/2020.06.19.20054809](http://dx.doi.org/10.1101/2020.06.19.20054809)[.](http://paperpile.com/b/DEg4PA/5cHl)
- 5. Lamy, P. *et al.* Paired Exome Analysis Reveals Clonal [Evolution](http://paperpile.com/b/DEg4PA/9Jxd) and Potential [Therapeutic](http://paperpile.com/b/DEg4PA/9Jxd) Targets in Urothelial Carcinoma. *Cancer Res.* **76**, 5894–5906 (2016).
- 6. Gerstung, M. *et al.* Reliable detection of subclonal [single-nucleotide](http://paperpile.com/b/DEg4PA/QfHnm) variants in tumour cell [populations.](http://paperpile.com/b/DEg4PA/QfHnm) *Nat. Commun.* **3**, 811 (2012).
- 7. Shukla, S. A. *et al.* Comprehensive analysis of [cancer-associated](http://paperpile.com/b/DEg4PA/n1zp) somatic mutations in class I HLA genes. *Nat. [Biotechnol.](http://paperpile.com/b/DEg4PA/n1zp)* **33**, 1152–1158 (2015).
- 8. Xie, C. *et al.* Fast and accurate HLA typing from short-read [next-generation](http://paperpile.com/b/DEg4PA/0uLm) sequence data with xHLA. *Proc. Natl. Acad. Sci. U. S. A.* **114**, [8059–8064](http://paperpile.com/b/DEg4PA/0uLm) (2017).
- 9. Szolek, A. *et al.* OptiType: precision HLA typing from [next-generation](http://paperpile.com/b/DEg4PA/Qizp) sequencing data. *[Bioinformatics](http://paperpile.com/b/DEg4PA/Qizp)* **30**, 3310–3316 (2014).
- 10. [Bjerregaard,](http://paperpile.com/b/DEg4PA/x5JU) A.-M., Nielsen, M., Hadrup, S. R., Szallasi, Z. & Eklund, A. C. MuPeXI: prediction of [neo-epitopes](http://paperpile.com/b/DEg4PA/x5JU) from tumor sequencing data. *Cancer Immunol. Immunother.* **66**, [1123–1130](http://paperpile.com/b/DEg4PA/x5JU) (2017).
- 11. Reynisson, B., Alvarez, B., Paul, S., Peters, B. & Nielsen, M. [NetMHCpan-4.1](http://paperpile.com/b/DEg4PA/OYvL) and [NetMHCIIpan-4.0:](http://paperpile.com/b/DEg4PA/OYvL) improved predictions of MHC antigen presentation by concurrent motif [deconvolution](http://paperpile.com/b/DEg4PA/OYvL) and integration of MS MHC eluted ligand data. *Nucleic Acids Res.* **48**, [W449–W454](http://paperpile.com/b/DEg4PA/OYvL) (2020).
- 12. Patro, R., Duggal, G., Love, M. I., Irizarry, R. A. & [Kingsford,](http://paperpile.com/b/DEg4PA/mPA8) C. Salmon provides fast and bias-aware [quantification](http://paperpile.com/b/DEg4PA/mPA8) of transcript expression. *Nat. Methods* **14**, 417–419 [\(2017\).](http://paperpile.com/b/DEg4PA/mPA8)
- 13. Davoli, T., Uno, H., Wooten, E. C. & Elledge, S. J. Tumor [aneuploidy](http://paperpile.com/b/DEg4PA/AQJ6i) correlates with markers of immune evasion and with reduced response to [immunotherapy.](http://paperpile.com/b/DEg4PA/AQJ6i) *Science* **355**, [\(2017\).](http://paperpile.com/b/DEg4PA/AQJ6i)
- 14. Danaher, P. *et al.* Gene expression markers of Tumor Infiltrating [Leukocytes.](http://paperpile.com/b/DEg4PA/HYTTn) *J [Immunother](http://paperpile.com/b/DEg4PA/HYTTn) Cancer* **5**, 18 (2017).
- 15. Rosenthal, R. *et al.* [Neoantigen-directed](http://paperpile.com/b/DEg4PA/rkR3Y) immune escape in lung cancer evolution. *Nature* **567**, [479–485](http://paperpile.com/b/DEg4PA/rkR3Y) (2019).
- 16. Manders, F. *et al.* [MutationalPatterns:](http://paperpile.com/b/DEg4PA/IAxc) the one stop shop for the analysis of mutational [processes.](http://paperpile.com/b/DEg4PA/IAxc) *BMC Genomics* **23**, 134 (2022).
- 17. Blokzijl, F., Janssen, R., van Boxtel, R. & Cuppen, E. [MutationalPatterns:](http://paperpile.com/b/DEg4PA/LcGG) [comprehensive](http://paperpile.com/b/DEg4PA/LcGG) genome-wide analysis of mutational processes. *Genome Med.* **10**, 33 [\(2018\).](http://paperpile.com/b/DEg4PA/LcGG)