Clonal Evolution of Chemotherapy-resistant Urothelial Carcinoma

Bishoy M. Faltas[1,2,3,#], **Davide Prandi**[4,#], **Scott T. Tagawa**[1,2,3], **Ana M. Molina**[1,2], **David M. Nanus**[1,2,3], **Cora Sternberg**[5], **Jonathan Rosenberg**[6], **Juan Miguel Mosquera**[7], **Brian Robinson**[7], **Olivier Elemento**[1,8,9], **Andrea Sboner**[1,7,9], **Himisha Beltran**[1,2,3,*], **Francesca Demichelis**[1,4,9,*], and **Mark A. Rubin**[1,3,7,*]

[1]Caryl and Israel Englander Institute for Precision Medicine, New York Presbyterian Hospital-Weill Cornell Medicine. New York, NY

[2]Department of Medicine, Division of Hematology and Medical Oncology, Weill Cornell Medicine. New York, NY

[3]Sandra and Edward Meyer Cancer Center at Weill Cornell Medicine. New York, NY

[4]Centre for Integrative Biology, University of Trento. Trento, Italy

[5]Department of Medical Oncology, San Camillo and Forlanini Hospitals. Rome, Italy

[6]Department of Medicine, Memorial Sloan Kettering Cancer Center. New York, NY

[7]Department of Pathology and Laboratory Medicine. Weill Cornell Medicine. New York, NY

[8]Department of Physiology and Biophysics. Weill Cornell Medicine. New York, NY

[9]Institute for Computational Biomedicine, Weill Cornell Medicine. New York, NY

Abstract

Chemotherapy-resistant urothelial carcinoma (UC) has no uniformly curative therapy. Understanding how selective pressure from chemotherapy directs UC's evolution and shapes its clonal architecture is a central biological question with clinical implications. To address this question, we performed whole-exome sequencing and clonality analysis of 72 UCs including 16 matched sets of primary and advanced tumors prospectively collected before and after chemotherapy. Our analysis provided several insights: (i) chemotherapy-treated UC is

characterized by intra-patient mutational heterogeneity and the majority of mutations are not shared, (ii) both branching evolution and metastatic spread are very early events in the natural history of UC; (iii) chemotherapy-treated UC is enriched with clonal mutations involving L1-cell adhesion molecule *(L1CAM)* and integrin signaling pathways; (iv) APOBEC induced-mutagenesis is clonally-enriched in chemotherapy-treated UC and continues to shape UC's evolution throughout its lifetime.

Introduction

Urothelial carcinoma (UC) results in 15,000 deaths annually in the United States[1]. Individuals with metastatic UC are standardly treated with platinum-based chemotherapy[2–5]. However, nearly all will progress and develop chemotherapy resistance[3,6,7]. Ultimately, the majority of patients will die of metastatic chemotherapy-resistant UC[2,3,5]. Little is known about the clonal architecture of advanced chemotherapy-treated UC or the evolutionary dynamics that lead to metastasis and chemotherapy resistance. Large genomic studies like The Cancer Genome Atlas (TCGA) have focused only on untreated primary tumors[8]. In particular, the extent to which chemotherapy-treated tumors share the genetic profile of the primary tumor remains unknown.

To understand the relative contributions of different subclones and the effects of chemotherapy as a selective pressure in UC, we performed whole exome sequencing and clonality analysis of matched sets of primary, metastatic and germline samples. Because cancers are genetically heterogeneous, we set out to address two fundamental questions: (i) what is the degree of clonal divergence between primary and metastatic UC? (ii) How does chemotherapy impact the genomic landscape of tumor cell populations in advanced and metastatic UC?

We employed the computational framework of CLONET (CLONality Estimate in Tumors) we developed previously[9] **(online methods)** to adjust genomic events for tumor purity and ploidy, and then determine the relative abundance of tumor cell subpopulations through clonality analysis of genomic lesions. By comparing the frequency and patterns of CLONET-adjusted events between primary and metastatic tumors obtained from different anatomical sites and at different time points over each patient's clinical course, we were able to reconstruct phylogenetic trees and compare clonal evolutionary patterns across the study cohort. We then proceeded to examine the clonally-enriched genomic signatures and trace the evolutionary footprints of mutagenesis mechanisms including the APOBEC3 family of cytidine-deaminases during each cancer's evolution.

Results

Clonal mutational heterogeneity in chemotherapy-treated UC

To characterize the clonal architecture of advanced chemotherapy-treated UC, we performed whole exome sequencing of 72 prospectively collected urothelial tumors from 32 patients including 16 matched sets of primary, metastatic UCs and germline samples and two rapid autopsy cases (Fig. 1, Supplementary Table 1). The study was designed to enrich for patients

with advanced disease, 28/32 (88%) of patients either presented with or developed metastatic disease during the study period (Fig. 1, Supplementary Table 1). Overall, the most frequent mutations and copy number alterations in our cohort were consistent with the results of the TCGA dataset of untreated UC[8] (Supplementary Fig. 1, Supplementary Table 2). We observed no statistically significant difference in the number of SNVs between pre-chemotherapy and post-chemotherapy tumors (Supplementary Fig. 2).

To compare the clonal structure of pre-chemotherapy and post-chemotherapy tumors across the study cohort, we investigated the number of private and shared mutations between the pre-chemotherapy and post-chemotherapy tumors within each patient as a fraction of the total mutational burden (Fig. 2a). On average, only 28.4% (range 0.2%–76.4%) of mutations were shared between pre- and post-chemotherapy samples (Fig. 2a). This effect was consistent across primary-primary tumor pairs and primary-metastatic tumor pairs (p=0.17, Wilcoxon test) (Fig. 2a). Surprisingly, even mutations in previously reported driver genes[10] including *PIK3CA*, *KMT2D (MLL2), ATM* and *TP53* were not consistently shared between matched pre-chemotherapy and post-chemotherapy tumors (Fig. 2b). We confirmed these findings with targeted sequencing of 250 common driver genes achieving an average coverage of 400x and an excellent concordance with variant allele frequencies obtained from whole exome sequencing (Pearson correlation = 0.93, $P<10^{-171}$) (Supplementary Fig. 3). Some post-chemotherapy tumors even evolved to develop different mutations in the same key gene. For example, in patient WCM077, the primary pre-chemotherapy tumor and the pelvic post-chemotherapy lymph node metastasis shared a *TP53* p.Y234C mutation while the post-chemotherapy lung metastasis had a separate private *TP53* p.G266V mutation that was not shared with the primary tumor (Fig. 2b, Supplementary Fig. 4). Taken together, our results demonstrate significant mutational heterogeneity in tumor samples from the same patient and suggest that chemotherapy is associated with a significant change in the mutational landscape of advanced urothelial carcinoma.

Early branching evolution in chemotherapy-treated UC

We conducted a phylogenetic analysis of 21 sets of matched tumors from patients from whom at least two tumor samples were available per patient using the parsimony ratchet method[11] (Fig. 3, Supplementary Fig. 5). This analysis revealed a pattern of early branching evolution with several successive waves of clonal expansion occurring early in each patient's UC. In every reconstructed evolutionary tree, the primary tumor was positioned as a branch indicating that the ancestral clone gave rise to multiple cell populations that evolved in parallel during the early stages of tumor evolution.

In order to better understand the pattern of clonal evolution during the course of chemotherapy, we followed another individual with UC from the time of diagnosis through death over a period of 16 months. We collected a total of 12 samples from 8 anatomical sites and at three different time-points: the primary untreated tumor obtained by transurethral resection of bladder tumor (TURBT) acquired at initial diagnosis, four areas of residual primary bladder tumor and one pelvic lymph node obtained by radical cystectomy and lymph node dissection following four cycles of gemcitabine-cisplatin chemotherapy, and six metastases obtained at time of rapid autopsy after docetaxel-ramucirumab therapy (4 distant

lymph nodes, 2 liver metastases) (Fig. 4a). We analyzed the genomes of these tumors to characterize the full evolutionary arc of the cancer from the time of diagnosis to death. Metastatic tumors obtained at autopsy harbored high clonal fractions (i.e. the percentages of clonal alterations compared to all alterations in a sample) (mean 83%, range 69%–91%) (Fig. 4b). Allele-specific copy number analysis identified a sub-clonal heterozygous deletion of the tumor suppressor *CDKN2A* in the primary tumor that evolved into a homozygous clonal deletion in the distant metastatic lymph nodes and liver lesions obtained from the autopsy. This was confirmed by fluorescence *in-situ* hybridization (FISH) (Fig. 4c) suggesting that *CDKN2A* loss was selected as the tumor evolved under pressure from chemotherapy. Comparison of clonality-adjusted frequencies of somatic alterations across all tumor samples from this patient revealed substantial heterogeneity. This heterogeneity followed several distinct patterns. Certain subclonal mutations (i.e. *RYR2, ANKRD62, NCOA3 and LSS*) were present in the primary tumor persisted and became enriched in the chemotherapy-treated metastatic lesions. Other subclonal mutations *(i.e. POLD2, FOXP1, FGFR4, TRRAP, and EGFR)* present in the primary were not observed in the metastatic lesions and were considered "private" to the primary tumor. We also observed other private mutations, which were present exclusively in the metastatic lesions (Fig. 4d). In fact, each tumor in this patient harbored a unique set of private mutations (mean 26.3, range 5–138) that was not shared with any of the other tumors (Fig. 4d).

Using clonally adjusted non-silent mutations from each tumor, we reconstructed the evolutionary tree of this patient's UC (**online methods**) (Fig. 4e). This reconstruction revealed a complex branching evolutionary pattern. Early truncal mutations (*RYR2, ANKRD62, NCOA3 and LSS*) were present in the initial founder clone and shared by all descendent clones. At each clonal divergence node, additional mutations were acquired including mutations in driver genes such as *TP53* and *TSC1*. Surprisingly, at the time of the patient's initial cancer diagnosis, at least 5 waves of clonal expansion (each represented by a branching node, numbered 1–5) (Fig. 4e) had already occurred from the lowest common ancestor as observed in the mutational analysis of the TURBT tumor. This pattern suggests that branching evolution was a very early event in this tumor's development. While the untreated TURBT tumor had a high fraction of tumor cells harboring founder mutations (Fig. 4e), it also had the farthest genetic distance from all other tumors. We attributed this long genomic distance to the high number of private mutations in the TURBT primary tumor sample (138 mutations) (Fig. 4d, Supplementary Table 3), many of which involve genes implicated in cellular responses to cisplatin including *POLD2* and *FOXP1*[12–15]. After neoadjuvant chemotherapy, these mutations disappeared from the evolutionary record when the cancer cells harboring them were likely eradicated by treatment and thus were not observed subsequently in any of the other tumors (Fig. 4d, 4e).

Interestingly, one of the earliest cancer cell clades had already separated at the first divergence node resulting in a population that metastasized to a pelvic lymph node (Fig. 4e). This lymph node was later dissected and removed at the time of radical cystectomy. Surgical removal of this lymph node possibly eliminated mutations from this particular clade from entering the genetic pool that later contributed to the development of additional distant metastases. On the other hand, by the time cystectomy took place, distant metastatic spread

had already occurred originating from a different cancer cell clade that branched at divergence node 5 to give rise to both bladder tumor #2 and all the lymph node and visceral distant metastases that were later collected at the time of the rapid autopsy. This transition from the primary to the metastatic state was marked by the acquisition of a non-silent mutation in tetraspanin8 (*TSPAN8*), a well-recognized pro-metastatic and angiogenesis-promoting gene[16–18](Fig. 4d, 4e). This sequence supports a possible role for this mutation as a key driver event underlying the metastatic spread of this patient's UC. Altogether, our data strongly suggests that both branching evolution and metastatic spread are very early events in the natural history of UC.

Heterogeneity in copy number alterations

To understand how copy number alterations (CNA) evolve throughout the lifetime of UC and during chemotherapy, we conducted a detailed analysis of somatic genomic aberrations. Hierarchical clustering of 44 tumor samples based on allele-specific copy number alterations (**online methods**) showed two distinct clusters (Fig. 5a). Cluster A was defined by 9p21 (*CDKN2A, CDKN2B and MTAP*) deletions in the setting of euploid copy number background. Cluster B was characterized by several enriched amplifications including 1q21.1 (*SETDB1 and MLLT11*) amplifications, (P=0.0002, Fisher's exact test) and 6p22.3 (*E2F3*) amplifications, (P=0.001, Fisher's exact test) (Supplementary Table 4). This cluster was also enriched with *TP53* mutations (P=0.0001, Fisher's exact test). The same CNAs clusters are consistently observed when extending our cohort with TCGA untreated UC[10] (Supplementary Fig. 6). We also observed an enrichment of tumors belonging to the TCGA bladder cancer cluster III ('basal/squamous-like')[10] in our copy number cluster A (Fisher's exact test P=0.02) (Supplementary Fig. 7). There was no statistically significant differential enrichment in the number of metastatic samples or chemotherapy treated-samples between the two clusters suggesting that these clusters may reflect a relatively stable feature of UC biology that is independent of treatment effects or disease stage. Overall, tumor samples from the same patient tended to cluster in the same group despite the presence of private CNAs. To quantify the degree of intra- and inter-patient heterogeneity, we interrogated CNAs from a panel of more that 30000 genes from the Ensembl catalog[19]. For each pair of tumor samples, we computed the Hamming Distance (HD) as the ratio between the number of genes that have different discrete copy number and the total number of genes analyzed. We identified a significant difference between intra-patient tumor pairs (median HD=0.20) and inter-patient pairs (median HD=0.53) (P=0.00000003, Wilcoxon test) (Fig. 5b). This limited intra-patient heterogeneity with respect to inter-patient heterogeneity suggests that each patient's cancer is relatively stable during evolution at the copy number level.

We investigated the frequency of combined copy number alterations and mutations constituting the *ATM/RB/FANCC* signature that was previously associated with chemotherapy response in the neoadjuvant setting[20]. We identified this signature in 11/15 (73.3%) in our pre-chemotherapy tumors and 11/29 (37.9%) (p=0.05) in post-chemotherapy tumors supporting the hypothesis that clones harboring these molecular alterations are likely to disappear after treatment and are superseded by tumor clones with wild-type *ATM/RB/FANCC* that eventually progress to metastatic chemotherapy-resistant disease.

Clonal enrichment of mutations in chemotherapy-treated UC

We hypothesized that the evolution of chemotherapy-treated UC would proceed in a direction that ultimately leads to the selection of mutations conferring proliferative or chemotherapy-resistance advantages. Using density analysis of CLONET-adjusted variant allele frequencies between pre-chemotherapy and post-chemotherapy tumors, we observed a significant increase in the number of clonal mutations in the post-chemotherapy samples across the study cohort (P=0.0134, Fisher's exact test) (Fig. 6a, Supplementary Fig. 8) confirming the association between chemotherapy and increased clonality. To dissect the functional impact of these clonally enriched mutations, we conducted gene set enrichment analysis (GSEA) to identify enriched pathways in post-chemotherapy samples (Fig. 6b). This analysis showed a clonal enrichment of mutations in pathways involved in the trans-membrane transport of small molecules (odds ratio = 1.9, FDR = 0.002) suggesting that mutations in multi-drug resistance genes may play a role in the progression of advanced chemotherapy-treated UC. In addition, GSEA demonstrated a significant enrichment in mutations mediating L1-cell adhesion molecule (*L1CAM*) (odds ratio = 1.9, FDR = 0.12), and integrin signaling pathways (odds ratio = 2.8, FDR = 0.02). The majority of mutations identified in the *L1CAM* and integrin signaling pathways (83% and 90% respectively) were missense mutations which can conceivably lead to gain-of-function molecular changes that activate these pathways. These results suggest a prominent role for mutations in *L1CAM* and integrin signaling pathways in conferring a selective advantage for resistance to chemotherapy in UC. Mutations in these pathways may also provide a potential mechanistic link between metastatic spread, tumor microenvironment, and drug-resistance that cooperate to promote tumor survival[21–25].

Mutagenesis mechanisms driving the evolution of UC

To characterize the evolution of mutational signatures in advanced chemotherapy-treated UC, we examined the six possible single-base substitutions (C>A, C>G, C>T, T>A, T>C, and T>G). We identified significant differences in these mutational patterns between chemotherapy-naïve and chemotherapy-treated tumors with a statistically significant enrichment of C>A and C>G changes in the chemotherapy-treated tumors (Fig. 7a).

To distinguish between potential mutagenic mechanisms responsible for these changes, we matched mutational patterns derived from statistical analysis of nucleotide changes to well-defined signatures of potential mutagens. We observed a significant increase in C>A nucleotide substitutions in tumors treated with cisplatin-based chemotherapy consistent with the specific mutagenesis signature induced in *C. elegans* genome after cisplatin treatment[26,27]. Further analysis of context motifs of various base substitutions showed enrichment in the (C> T or G changes at the TCW motifs, W–A or T) (Fig. 7b) which is highly suggestive of APOBEC-induced mutagenesis[28,29]. To confirm this finding, we compared the signature in our cohort to previously reported Sanger signatures[28–30] (**online methods**). We observed 4 distinct signatures in our cohort (Fig. 7c). The first signature was very similar to Sanger signatures 2 and 13, attributed to APOBEC mutagenesis[28–31]. We detected three additional signatures corresponding to previously described mutagenic processes associated with age, smoking and *ERCC2* mutations[28,30]. The low frequency of *ERCC2* mutations in our cohort of chemotherapy-treated UC (Supplementary Fig. 1) is

consistent with previous reports suggesting that *ERCC2* mutations are enriched in responders to cisplatin-based chemotherapy[32,33] and likely selected against in tumors that progress through chemotherapy.

Because of the prominence of APOBEC-induced mutagenesis in UC, we focused on understanding how APOBEC-induced mutations evolve during chemotherapy by comparing the frequency of APOBEC-induced mutations in chemotherapy-naïve and chemotherapy-treated tumors. We observed a significant enrichment in the APOBEC3-induced mutagenesis (C > T or G changes at the TCW motifs, W–A or T) in post-chemotherapy tumors (Fig. 7d). To dissect the relative contributions of individual members of APOBEC3 cytosine-deaminases to this enrichment, we examined the preferred motif contexts favored by individual APOBEC enzymes for mutating respective cytosines[34,35]. APOBEC3A favors YTCA, while APOBEC3B favors RTCA motifs (wherein Y = pyrimidine, R = purine)[34]. APOBEC3G induces cytosine substitutions in the single stranded DNA overhang strand with a preference for the 5′-CCC-3′ motifs[35,36]. APOBEC3F preferentially mutates cytosines in the TTC motif (where the underlined C is the mutated nucleotide)[37]. We detected a significant enrichment in APOBEC3A-induced mutations (P=0.00001, Fisher's exact test), and a similar enrichment of APOBEC3B mutagenesis (P=0.0395, Fisher's exact test) in post-chemotherapy tumors. In contrast, APOBEC3G mutagenesis was substantially decreased in post-chemotherapy tumors (Fig. 7d). Furthermore, we observed a corresponding statistically significant increase in the clonality of APOBEC-induced mutations in post-chemotherapy tumors (Fig. 7e). Enrichment analysis of APOBEC-induced mutations highlighted key pathways involved in chemotherapy resistance including the ABC family of proteins (odds ratio=2.7, P=0.038, Fisher's exact test) and homologous recombination DNA-damage repair (odds ratio=3.8, P=0.033, Fisher's exact test) (Supplementary Table 5). Our findings suggest that the APOBEC mutational process is not merely a transient event in early UC oncogenesis but that it continues to shape the evolution of advanced UC and may promote clonal expansions of chemotherapy-resistant clones.

Discussion

Advanced chemotherapy-resistant UC remains a formidable clinical challenge with limited therapeutic options[38]. Whole-exome analysis of matched samples from the same patient from different anatomical sites and at sequential time points offers a unique opportunity to reconstruct the evolutionary dynamics and understand the mutagenic pressures shaping the evolution of primary untreated UC to advanced chemotherapy-treated UC.

Our analyses identified substantial spatial and temporal heterogeneity between tumors separated in time or by anatomical location within the same patient. The majority of mutations in the post-chemotherapy tumors were not shared with primary chemotherapy-naive tumors. Branching evolution was the predominant path from primary chemotherapy-naïve UC to advanced chemotherapy-treated UC. Very early in this path, several clonal waves separate from the original founder clone, many of which metastasize early in the tumor's lifetime and continue to evolve in parallel with the primary tumor. Our findings shed light on the importance of addressing gaps in the existing knowledge of the clonality of early events in UC oncogenesis including multicentricity and malignant seeding, which could lead

to alternative interpretations of the phylogenetic trees. Additionally, our findings suggest that extensive heterogeneity and early branching evolution should be taken into consideration as additional layers of biological complexity that go beyond the traditional two-pathway UC oncogenesis model and potentially eclipse grade and stage classifciations[39].

We demonstrate for the first time that chemotherapy-treated UC is significantly clonally enriched in mutations in *L1CAM* and integrin-signaling pathways. The majority of these mutations were missense mutations that could potentially lead to activation of these pathways but the precise functional impact of these mutations warrants future studies. Our results are consistent with data in pre-clinical models of other tumor types such as cholangiocarcinoma, ovarian carcinoma and pancreatic ductal adenocarcinoma demonstrating that *L1CAM* plays a key role in cisplatin-resistance and protecting cells from apoptosis[40–43]. *L1CAM* directly binds to integrin receptors via its RGD-motif in the sixth Ig-domain[44,45] and there is considerable cross-talk between the two pathways[46,47]. Previous studies demonstrated that integrin-signaling plays important roles in overriding chemotherapy-induced cell cycle arrest and apoptosis in small-cell lung cancer cells through activation of the PI3-kinase pathway[48,49]. Overexpression of Beta 1-integrin in hepatocellular carcinoma cell lines protected them against apoptosis induced by chemotherapeutic agents by activating MAP-kinase signaling[50]. Stimulating Beta1-integrin with an antibody ligand in leukemia cells prevented procaspase-8-mediated induction of apoptosis in a PI3K-dependent manner[51]. Collectively, these observations suggest that alterations in *L1CAM* and integrin signaling pathways potentially play a key role in chemotherapy-resistance in UC and provide a mechanistic intersection between the microenvironment and drug-resistance through cell-adhesion-mediated drug resistance (CAM-DR) phenomenon[21,52,53]. This phenomenon has been implicated in chemotherapy resistance in several malignancies[54,55]. *L1CAM* is potentially targetable with antibodies that have demonstrated efficacy in xenograft animal models of cholangiocarcinomas[56,57], ovarian and pancreatic ductal carcinomas[58]. Focal Adhesion Kinase (FAK) inhibitors, which target integrin signaling, have been shown to profoundly sensitize cancer cells to chemotherapy and novel molecular therapeutics and are currently in early phase clinical trials[59,60]. Our results suggest that similar therapeutic approaches merit further study in chemotherapy-resistant UC.

We clearly demonstrate an increase in APOBEC signatures in chemotherapy-treated tumors. One possible explanation for this interesting finding is that platinum-based chemotherapy increases the formation of APOBEC mutagenesis-prone single-stranded DNA (ssDNA)[61,62]. This ssDNA is formed during $5' \rightarrow 3'$ resection that occurs at DNA double-strand-breaks during homology-directed repair [63]. While allowing for error-free repair of these double-strand-breaks DNA induced by the excision of platin-DNA adducts this process may potentially increase the availability of intermediary ssDNA to APOBEC mutagenesis[63]. Our results also suggest that APOBEC3A is the main enzyme responsible for mutagenesis in advanced chemotherapy-treated UC. This is in accordance with recent data suggesting that APOBEC3A-mediated mutagenesis is the key mutagenic cytidine deaminase in most tumor types because of its high proficiency in generating DNA breaks[64]. These findings suggest a potential mechanism by which chemotherapy acts to increase the genomic diversity of chemotherapy-treated tumors that requires future study. Our data demonstrate that the clonal

evolution of chemotherapy-treated UC is characterized by a dramatic divergence of the mutational landscape in the face of relative stability at the copy-number level over the lifetime of each tumor. This finding potentially reflects the dominance of APOBEC-induced mutagenesis in UC as a mechanism that preferentially induces single nucleotide changes throughout the tumors' lifetime and during chemotherapy. An alternative explanation for the marked genomic alterations we observed in chemotherapy-treated samples is that a certain degree of genetic drift occurs over time irrespective of the effect of chemotherapy. However, it is unlikely that genetic drift is the sole mechanism accounting for the genetic heterogeneity we observed because chemotherapy is a potent selective pressure that is expected to alter the evolutionary dynamics affecting the pace and steering the direction of genetic drift. In fact, our results are consistent with evolutionary models suggesting that cancer's adaptation ensues from the interaction between stochastic processes such as mutation generation and clonal selection, which is a deterministic phenomenon[65]. Our findings support this evolutionary model by demonstrating a complex and dynamic interplay between mutagenic mechanisms such as APOBEC-induced mutagenesis, and extrinsic selective pressures such as chemotherapy to constantly shape the clonal evolution of UC. As genetic information is passed from parent to progeny clones, each process leaves an evolutionary record of molecular alterations in descendent clones that allows reconstruction of the process. However, it is important to note that this record fully exists only in clones that survive selection. Unfit clones are eliminated from the record and can only be captured by serial sampling of cancer cells throughout the tumor's lifetime whereas resistant clones are selected to expand and supersede previous clonal waves.

One major strength of our study is that it is the most comprehensive study of the clonal evolution of chemotherapy-resistant urothelial carcinoma. Limitations of our study include a small sample size. Of note, we included muscle invasive tumor samples from patients who were never treated with chemotherapy as controls.

Our findings have several potential clinical implications: First, genomic divergence between untreated and treated clones suggests that clinically actionable molecular targets in metastatic chemotherapy-treated tumors may be missed when relying only on biopsies of untreated primary tumors at the time of diagnosis, and that repeat metastatic biopsies during the course of clinical care would be needed to detect the most recent version of the rapidly changing molecular landscape of a given patient's UC. Second, further study of the functional role of *L1CAM* and integrin-signaling in mediating chemotherapy-resistance in UC could lead to a potential strategy for reversing or preventing chemotherapy resistance by targeting these pathways. Third: despite its initial effectiveness in eliminating cancer cells, platinum-based chemotherapy is associated with unintended significant mutagenic editing of the genomic landscape of post-chemotherapy tumors. Our insight into the nature of these edits is crucial towards a complete understanding of the basis of chemotherapy-resistance in advanced UC, which may lay the foundation for the development of rational therapeutic strategies for preventing the emergence or reversing the chemotherapy-resistant state of UC.

In summary, our results demonstrate that advanced chemotherapy-treated UC undergoes extensive and dynamic clonal evolution throughout the lifetime of the tumor with significant genetic editing that continues during and after chemotherapy. Our findings lay the

foundation for an evolutionary understanding of advanced chemotherapy-treated UC and present opportunities for advancing cancer precision medicine.

Online Methods

Patient enrollment and tumor procurement

All experimental procedures were carried out in accordance with approved guidelines and were approved by the Institutional Review Board (IRB) at Weill-Cornell Medicine. Patients signed informed consent under (IRB) approved protocol (IRB #1305013903). Clinical information was collected from the chart. Smoking status was collected from self-administered questionnaires. Tumor samples were obtained from patients through surgical resection or core biopsies.

Rapid autopsy procedures

The Englander Institute for Precision Medicine at Weill Cornell Medicine-New York Presbyterian has been established to promote personalized medicine focused on molecular diagnostics and therapeutics. Two patients in our series selected the option to be enrolled in the IRB-approved rapid autopsy program. In addition, patients' next-of-kin provided written consent before autopsy. The WCM117 and WCM259 rapid autopsies were conducted within 6 hours after death. A systematic autopsy protocol is followed where normal and malignant fresh tissue is collected, allocating samples to be snap frozen or formalin-fixed. The goal is to maximize the amount of tissue collected for research purposes. Once the tissue harvest is complete, the autopsy proceeds in accordance with the protocol established by the WCM Autopsy Service. For our current study, tissue samples from multiple sites were procured from each patient as detailed above. After H&E evaluation and frozen slide annotation, DNA was extracted for WES.

DNA extraction and next generation sequencing

In this study, we used a New York State approved whole exome sequencing assay developed in our CLIA laboratory called, EXaCT-1[1]. After macro-dissection of target lesions, tumor DNA was extracted from FFPE or cored OCT-cryopreserved tumors using the Promega Maxwell 16 MDx (Promega, Madison, WI, USA). Germline DNA was extracted from peripheral blood mononuclear cells using the same method. Pathological review by one of the study pathologists (JMM, BR, MAR) confirmed the diagnosis and determined tumor content. A minimum of 200ng of DNA was used for whole exome sequencing. DNA quality was determined by TapeStation Instrument (Agilent Technologies, Santa Clara, CA) and was confirmed by real-time PCR before sequencing. Sequencing was performed using Illumina HiSeq 2500 (2x100bp). A total of 21,522 genes were analyzed with an average coverage of 85x (range 60–102, Supplementary Table 6) using Agilent HaloPlex Exome (Agilent Technologies, Santa Clara, CA). We developed a targeted sequencing assay of 250 cancer genes (*referred to as N250*) using hybrid-capture SeqCap EZ Choice Enrichment Kits (Roche Sequencing, Pleasanton, CA) (Supplementary Table 7 and Supplementary Table 8). Sequencing was performed using Illumina HiSeq 2500 (PE 2x75) achieving an average coverage of 400x.

Study sample size definition

Our study includes 72 samples from 32 patients. However, patient WCM117 comprises 12 samples (17% of the total) and to avoid possible statistical biases in the analysis, we utilized WCM117 (12 samples) only in Fig. 4, as a specific case study to better understand how chemotherapy shapes evolution. Finer analysis, as allele specific copy number (Fig. 5a) and SNVs clonality (Fig. 6a) also require an estimate of ploidy and purity. After manual inspection of CLONET outputs, we ended up with 44 samples in 25 patients with reliable ploidy and purity estimates. The exact number of samples and patients used in each figure is reported in Supplementary Notes.

Sequence data processing pipeline

All the study samples data were processed through the computational analysis pipeline of the Institute of Precision Medicine at Weill Cornell/New York Presbyterian Hospital (IPM-Exome-pipeline)[1]. Raw reads quality has been assessed with FASTQC as described previously[1]. Multi-sample patient data were tested for genotype distance using SPIA[2]. Pipeline output includes segment DNA copy number data, somatic copy-number aberrations (SCNAs) (Supplementary Table 9), and putative somatic single nucleotide variants (SNVs) (Supplementary Table 10). Finally, to assess tumor ploidy and purity we applied CLONET[3] on segmented data and allelic fraction (AF) of germline heterozygous SNP loci (termed *informative SNPs,* see **Allele specific copy number analysis**). Upon visual inspection of CLONET output, 53 of 72 samples data were deemed appropriate for downstream copy number analysis; excluded samples data were not associated with chemotherapy treatment (p=0.4384) or biopsy site (p=1).

We have previously published on how EXaCT-1 was developed and optimized for use with FFPE and Frozen samples- in contrast to many research-grade assays[1]. To further ensure that our results reflect biological effects rather than technical variability between different sample types, we took special care to account for variations in tumor content for each sample in order to correctly map the clonal evolution of UC. In particular, we did not observe any significant difference (P=0.1, Wilcoxon test) when comparing FFPE and fresh samples purity (Supplementary Fig. 9a). Similarly, we did not detect differences (P=0.137, Wilcoxon test) in the numbers of identified non-silent SNVs between FFPE and fresh samples (Supplementary Fig. 9b).

Allele-specific copy number analysis

Somatic copy number altered regions are defined by the log2 of the ratio between the tumor and normal local coverage normalized by the global tumor and normal coverage ratio (named *log2R*). CLONET refines copy number data adjusting each *log2R* to account for both aneuploidy and tumor purity. Combining purified *log2R* values and AF of *informative SNPs*, CLONET assigns allele-specific copy number values, represented as a pair (*cnA*, *cnB*), to each genomic segment[4]. Quality filters require at least 10 informative SNPs and a mean coverage of 20 reads to call allele-specific values of a segment. If a segment does not pass filters, adjusted *log2R* values below −0.4 (above 0.4) were categorized as copy number loss (gain).

Differential copy number analysis between pre- and post-chemotherapy has been performed on a set of 1160 genes selected among putative cancer genes [COSMIC[5] and Intogen[6]] (*cancerGene*s) and bladder-specific genes (Supplementary Table 11).

Unsupervised clustering analysis was performed on allele-specific copy number of *cancer Genes* by means of hierarchical clustering. Briefly, each gene *gID* is represented as a pair of real values corresponding to the allele-specific copy number of the genomic segment comprising *gID*. Then, the Euclidean distance on allele-specific copy number calls is used as distance; this approach distinguishes between ambiguous cases such as copy number wild type status (allele-specific values (1,1)) and copy number neutral loss (2,0) both corresponding to *log2R*=0 leading to more informative clusters (Supplementary Table 4). Copy number based analysis identified two clusters that we named WCM_A and WCM_B (Fig. 5, Supplementary Fig. 6), which we compared with the original four TCGA clusters (I, II, III, IV) resulted from integrated analysis of mRNA, miRNA and protein data[7] (Supplementary Fig. 7). Given a TCGA cluster X, we tested the null hypothesis that clusters WCM_A and WCM_B contain the same proportion of samples from X using Fisher's exact test (significance level = 0.05).

Single nucleotide variations analysis

To improve the quality of SNVs calls in targeted exons, we applied an integrated approach. We first ran both MuTect[8] and SNVseeqer[9] to nominate putative aberrant genomic positions. Then, closely looked at the identified positions by means of ASEQ[10] in normal and germline samples executing pileup analysis; for each single nucleotide position identified as putative aberrant, ASEQ returns information about the read count for each of the 4 bases A, C, G, and T. To reduce false positives, we required base quality and read quality above 20. Finally, a genomic position is considered aberrant in a tumor sample, if the read count of the alternative base is 0 in the matched germline and 3 or more in the tumor. This step allows to (i) filter out remaining germline SNPs, that is, positions where the alternative base is present in the control sample; (ii) check for the presence of SNVs with low AFs in patient's multiple samples data. This step is particularly relevant in this study, because SNVs identification methods are designed to work on a single normal-tumor pair, and they do not consider that samples from the same patient could share the same SNVs because they are not independent samples. Finally, we annotated genomic position with information relevant for cancer analysis with Oncotator[11] (v1.8.0.0). We exploited the last Oncotator datasource corpus including annotations about gene/transcript names, functional consequence (e.g. missense or nonsense), the predicted impact on protein function, annotation from cancer specific resources as COSMIC or TCGAscape, and possible published results about the specific mutation (Supplementary Table 10). A full description of the resources used by Oncotator is available in the tool help page. We identified 13 possible functional consequences (Supplementary Fig. 10) described in the Sequence Ontology[12]. To avoid overestimating divergence among samples from the same patient, we were conservative in defining SNVs that more likely produce a change in the protein, i.e., they affect the phenotype, and we considered non-silent only mutation classified as missense, nonsense, splice_site, nonstop, and start codon. However, for the analysis described in Fig. 7, we included all the identified mutations (Supplementary Fig. 10) as mutational mechanisms also affect mutations with

neutral functional effects. To further confirm our findings and to check for possible biases introduced by the definition of silent mutations, we also performed the analysis of Figs. 2, 3, 4, and **6** using both silent and non-silent SNVs obtaining comparable results (Supplementary Fig. 11, 12, 13 and 14).

SNVs enrichment analysis

Gene Set Enrichment Analysis (GSEA) refers to a computational method that identifies set of genes that are statistically enriched for a given observable variable. In this study, we interrogated the REACTOME pathway database to search for pathways showing a significant increase in the number of SNVs. We applied Fisher's exact test followed by FDR correction by Benjamin-Hochberg (BH) procedure. REACTOME pathways with FDR 0.2 are reported in Fig. 6b. We also checked if, among significant pathways, there is statistically significant difference between pre- and post-chemotherapy samples. Given a pathway P, Fisher's exact test determines the probability that the number of SNVs in P is different when considering pre- and post-chemo samples. After FDR correction (BH procedure), we highlighted nodes in Fig.6b with FDR 0.2.

Phylogenetic analysis

High-quality variants identified in the previous steps were used to re-construct the phylogenetic tree of each patient using the parsimony Ratchet method[13] (Fig. 3, Fig. 4e and Supplementary Fig. 5). In this representation, each node models a population of tumor cells. Nodes with no children, named leafs, represent cell populations from a tumor sampling, i.e., a tumor biopsy. Internal nodes model inferred tumor cell populations from observed SNVs. The node named WT represents a hypothetical population of wild type cells (cells with no somatic aberrations). In phylogenetic trees, an edge connects two nodes; the length of an edge is proportional to the number of SNVs. For instance, in Fig. 4e, node 1 corresponds to the least common ancestor inferred from all the available biopsies with the number of mutations proportional to the length of the edge from WT to 1. Node 1 is also connected to the Pelvic LN met and to the inferred cell population 2. A branch represents a time point in the evolution of the tumor where two distinct cell populations emerge; the length of the branches models the number of SNVs that are private to each population. In Fig. 4e, we observed few private mutations in the Pelvic LN met with respect to the number of SNVs shared by all the other samples, as supported by heatmap in Fig. 4d.

Clonality of single nucleotide variants

Original CLONET implementation allows for the computation of the clonality of an SNV with copy number normal genomic segment, i.e. the segment has an allele-specific copy number (1, 1). Here we extended it to allow for SNV clonality estimation independently from the copy number status of the genomic segments in which the SNV lies. Given the tumor purity *P*, the allele specific copy number (*cnA*, *cnB*), and the number of reads *nRef* and *nAlt* supporting the reference and the alternative base, respectively, estimating the clonality of a SNV requires to compute expected AF. We observed that AF could assume only a finite number of values given the DNA copy number state. For instance, let's assume that a locus is aberrant (mutated) in one allele and wild-type (not mutated) in two alleles, i.e.

SNV in a copy number aberrant segment (CN=3). In this case, with tumor cellularity equal to 100% and clonal SNV, the VAF would be equal to 1/3. Given the number of allele *cnSNV* harboring a *SNV*, the expected VAF is defined as

$$exp\,VAF = \frac{cnSNV * P}{2 * (1-P) + (cnA + cnB) * P}$$

To estimate *cnSNV*, we followed parsimony approach assuming *cnSNV* that best explains VAF adjusted by DNA admixture, computed as previously described[3]. We computed the clonality of *SNVs* using the distance between the observed and expected VAF as

$$Cl = \frac{(VAF - exp\,VAF) * (2 * (1-P) + (cnA + cnB) * P)}{cnSNV * P}$$

Complete proof in Supplementary Notes.

Single nucleotide variants signatures

SNVs are partitioned into six mutation classes (column "Mutation class", Supplementary Table 10) corresponding to six types of base pair substitution, C>A, C>G, C>T, T>A, T>C, T>G. The null hypothesis that pre- and post-chemotherapy samples are equally likely to harbor SNVs of a specific mutation class is tested with Fisher's exact test (Fig. 7a). The fingerprint of a SNV includes the two bases immediately 5′ and 3′ to each SNV position (column "Genomic context", Supplementary Table 10) for a total of 96 possible mutation fingerprints[14]. Fisher's exact test adjusted for multiple hypotheses testing with Benjamin-Hochberg procedure returns the likelihood that a mutation fingerprint is enriched in pre- or post-chemotherapy samples (Fig. 7b). As the set of mutation fingerprints of a tumor sample is a proxy for the mutational processes that shape the cancer genome, we studied the mutational signatures of our study samples and compared them the with the Sanger signatures[14] applying the same approach recently proposed[15]. Briefly, the Sanger signatures were obtained from the identification of 30 mutational processes signatures (named *Sanger signatures*) upon application of the original tool on more than 10,000 samples from 40 distinct human cancer types[14]. In our dataset, we identify 4 mutational signatures (Fig. 7c); the Sanger signature analysis reveals that APOBEC proteins play a role in the mutational processes shaping UC genomes. We checked for statistically significant differences between pre- and post-chemotherapy of individual members of the APOBEC family using Fisher's exact test followed by Benjamin-Hochberg FDR correction (Fig. 7d). To test if the clonality of APOBEC induced SNVs is enriched in post-chemotherapy samples, we dichotomized the SNV clonality levels (threshold 0.6) and then we applied Fisher's exact test (Fig. 7e).

Fluorescence *in situ* hybridization

Two 4-μm-thick tissue sections from each block were cut for FISH analysis. *CDKN2A* deletion was determined using FISH probe (BAC clone RP11-149I2) and a reference probe,

located at 9p21. At least 100 nuclei were evaluated per sample using a fluorescence microscope (Olympus BX51; Olympus Optical).

Statistical analysis

For statistical tests, two-sided Mann–Whitney–Wilcoxon test (referred to as Wilcoxon test in the main text) was used to check for significant differences between two distributions. The two-sided Fisher's exact test was applied to determine whether the deviations between the observed and the expected counts were significant. When appropriate p-values were adjusted for multiple hypotheses testing with Benjamin-Hochberg procedure. Boxplot statistics were computed with the function "boxplot" of R programming language. No statistical methods were used to predetermine study sample size.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

References

1. Howlader, N., et al. SEER cancer statistics review, 1975–2010. Bethesda, MD: National Cancer Institute; 2013.

2. von der Maase H, et al. Long-term survival results of a randomized trial comparing gemcitabine plus cisplatin, with methotrexate, vinblastine, doxorubicin, plus cisplatin in patients with bladder cancer. J Clin Oncol. 2005; 23:4602–8. [PubMed: 16034041]

3. Clark PE, et al. Bladder cancer. J Natl Compr Canc Netw. 2013; 11:446–75. [PubMed: 23584347]

4. Sternberg CN, et al. Randomized phase III trial of high-dose-intensity methotrexate, vinblastine, doxorubicin, and cisplatin (MVAC) chemotherapy and recombinant human granulocyte colony-stimulating factor versus classic MVAC in advanced urothelial tract tumors: European Organization for Research and Treatment of Cancer Protocol no. 30924. J Clin Oncol. 2001; 19:2638–46. [PubMed: 11352955]

5. Sternberg CN, et al. Seven year update of an EORTC phase III trial of high-dose intensity M-VAC chemotherapy and G-CSF versus classic M-VAC in advanced urothelial tract tumours. Eur J Cancer. 2006; 42:50–4. [PubMed: 16330205]

6. Bellmunt J, et al. Phase III trial of vinflunine plus best supportive care compared with best supportive care alone after a platinum-containing regimen in patients with advanced transitional cell carcinoma of the urothelial tract. J Clin Oncol. 2009; 27:4454–61. [PubMed: 19687335]

7. Bellmunt, JaPDP. New therapeutic challenges in advanced bladder cancer. Seminars in oncology. 2012; 39:598–607. [PubMed: 23040256]

8. Comprehensive molecular characterization of urothelial bladder carcinoma. Nature. 2014; 507:315–22. [PubMed: 24476821]

9. Prandi D, et al. Unraveling the clonal hierarchy of somatic genomic aberrations. Genome Biol. 2014; 15:439. [PubMed: 25160065]

10. Network CGAR. Comprehensive molecular characterization of urothelial bladder carcinoma. Nature. 2014; 507:315–22. [PubMed: 24476821]

11. Nixon KC. The Parsimony Ratchet, a New Method for Rapid Parsimony Analysis. Cladistics. 1999; 15:407–414.

12. Zheng Y, et al. Identification of Pold2 as a novel interaction partner of protein inhibitor of activated STAT2. Int J Mol Med. 2012; 30:884–8. [PubMed: 22824807]

13. Lee YS, Gregory MT, Yang W. Human Pol ζ purified with accessory subunits is active in translesion DNA synthesis and complements Pol η in cisplatin bypass. Proc Natl Acad Sci U S A. 2014; 111:2954–9. [PubMed: 24449906]

14. van Boxtel R, et al. FOXP1 acts through a negative feedback loop to suppress FOXO-induced apoptosis. Cell Death Differ. 2013; 20:1219–29. [PubMed: 23832113]

15. Choi EJ, et al. FOXP1 functions as an oncogene in promoting cancer stem cell-like characteristics in ovarian cancer cells. Oncotarget. 2016; 7:3506–19. [PubMed: 26654944]

16. Greco C, et al. E-cadherin/p120-catenin and tetraspanin Co-029 cooperate for cell motility control in human colon carcinoma. Cancer Res. 2010; 70:7674–83. [PubMed: 20858717]

17. Wei L, Li Y, Suo Z. TSPAN8 promotes gastric cancer growth and metastasis via ERK MAPK pathway. Int J Clin Exp Med. 2015; 8:8599–607. [PubMed: 26309511]

18. Zöller M. Tetraspanins: push and pull in suppressing and promoting metastasis. Nat Rev Cancer. 2009; 9:40–55. [PubMed: 19078974]

19. Kinsella RJ, et al. Ensembl BioMarts: a hub for data retrieval across taxonomic space. Database (Oxford). 2011; 2011:bar030. [PubMed: 21785142]

20. Plimack ER, et al. Defects in DNA Repair Genes Predict Response to Neoadjuvant Cisplatin-based Chemotherapy in Muscle-invasive Bladder Cancer. Eur Urol. 2015; 68:959–67. [PubMed: 26238431]

21. Eke I, Cordes N. Focal adhesion signaling and therapy resistance in cancer. Semin Cancer Biol. 2015; 31:65–75. [PubMed: 25117005]

22. Maiuthed A, Chanvorachote P. Cisplatin at sub-toxic levels mediates integrin switch in lung cancer cells. Anticancer Res. 2014; 34:7111–7. [PubMed: 25503138]

23. Seguin L, Desgrosellier JS, Weis SM, Cheresh DA. Integrins and cancer: regulators of cancer stemness, metastasis, and drug resistance. Trends Cell Biol. 2015; 25:234–40. [PubMed: 25572304]

24. Tsutsumi S, et al. L1 Cell adhesion molecule (L1CAM) expression at the cancer invasive front is a novel prognostic marker of pancreatic ductal adenocarcinoma. J Surg Oncol. 2011; 103:669–73. [PubMed: 21360711]

25. Weidle UH, Eggle D, Klostermann S. L1-CAM as a target for treatment of cancer with monoclonal antibodies. Anticancer Res. 2009; 29:4919–31. [PubMed: 20044598]

26. Meier B, et al. C. elegans whole-genome sequencing reveals mutational signatures related to carcinogens and DNA repair deficiency. Genome Res. 2014; 24:1624–36. [PubMed: 25030888]

27. Szikriszt B, et al. A comprehensive survey of the mutagenic impact of common cancer cytotoxics. Genome Biol. 2016; 17:99. [PubMed: 27161042]

28. Alexandrov LB, et al. Signatures of mutational processes in human cancer. Nature. 2013; 500:415–21. [PubMed: 23945592]

29. Nik-Zainal S, et al. Mutational processes molding the genomes of 21 breast cancers. Cell. 2012; 149:979–93. [PubMed: 22608084]

30. Kim J, et al. Somatic ERCC2 mutations are associated with a distinct genomic signature in urothelial tumors. Nat Genet. 2016

31. Roberts SA, et al. An APOBEC cytidine deaminase mutagenesis pattern is widespread in human cancers. Nat Genet. 2013; 45:970–6. [PubMed: 23852170]

32. Liu D, et al. Clinical Validation of Chemotherapy Response Biomarker ERCC2 in Muscle-Invasive Urothelial Bladder Carcinoma. JAMA Oncol. 2016

33. Van Allen EM, et al. Somatic ERCC2 mutations correlate with cisplatin sensitivity in muscle-invasive urothelial carcinoma. Cancer Discov. 2014; 4:1140–53. [PubMed: 25096233]

34. Hoopes JI, et al. APOBEC3A and APOBEC3B Preferentially Deaminate the Lagging Strand Template during DNA Replication. Cell Rep. 2016; 14:1273–82. [PubMed: 26832400]

35. Lackey L, et al. APOBEC3B and AID have similar nuclear import mechanisms. J Mol Biol. 2012; 419:301–14. [PubMed: 22446380]

36. Chelico L, Pham P, Goodman MF. Stochastic properties of processive cytidine DNA deaminases AID and APOBEC3G. Philos Trans R Soc Lond B Biol Sci. 2009; 364:583–93. [PubMed: 19022738]

37. Harris RS, Liddament MT. Retroviral restriction by APOBEC proteins. Nat Rev Immunol. 2004; 4:868–77. [PubMed: 15516966]

38. Faltas BM, Karir BS, Tagawa ST, Rosenberg JE. Novel molecular targets for urothelial carcinoma. Expert Opin Ther Targets. 2015; 19:515–25. [PubMed: 25633079]

39. Knowles MA, Hurst CD. Molecular biology of bladder cancer: new insights into pathogenesis and clinical diversity. Nat Rev Cancer. 2015; 15:25–41. [PubMed: 25533674]

40. Yoon H, et al. L1 cell adhesion molecule and epidermal growth factor receptor activation confer cisplatin resistance in intrahepatic cholangiocarcinoma cells. Cancer Lett. 2012; 316:70–6. [PubMed: 22088438]

41. Stoeck A, et al. L1-CAM in a membrane-bound or soluble form augments protection from apoptosis in ovarian carcinoma cells. Gynecol Oncol. 2007; 104:461–9. [PubMed: 17030349]

42. Sebens Müerköster S, et al. Drug-induced expression of the cellular adhesion molecule L1CAM confers anti-apoptotic protection and chemoresistance in pancreatic ductal adenocarcinoma cells. Oncogene. 2007; 26:2759–68. [PubMed: 17086212]

43. Sebens Müerköster S, et al. alpha5-integrin is crucial for L1CAM-mediated chemoresistance in pancreatic adenocarcinoma. Int J Oncol. 2009; 34:243–53. [PubMed: 19082495]

44. Kiefel H, et al. L1CAM: a major driver for tumor cell invasion and motility. Cell Adh Migr. 2012; 6:374–84. [PubMed: 22796939]

45. Gast D, et al. The RGD integrin binding site in human L1-CAM is important for nuclear signaling. Exp Cell Res. 2008; 314:2411–8. [PubMed: 18555990]

46. Voura EB, Ramjeesingh RA, Montgomery AM, Siu CH. Involvement of integrin alpha(v)beta(3) and cell adhesion molecule L1 in transendothelial migration of melanoma cells. Mol Biol Cell. 2001; 12:2699–710. [PubMed: 11553709]

47. Burgett ME, et al. Direct contact with perivascular tumor cells enhances integrin $\alpha v \beta 3$ signaling and migration of endothelial cells. Oncotarget. 2016

48. Hodkinson PS, et al. ECM overrides DNA damage-induced cell cycle arrest and apoptosis in small-cell lung cancer cells through beta1 integrin-dependent activation of PI3-kinase. Cell Death Differ. 2006; 13:1776–88. [PubMed: 16410797]

49. Sethi T, et al. Extracellular matrix proteins protect small cell lung cancer cells against apoptosis: a mechanism for small cell lung cancer growth and drug resistance in vivo. Nat Med. 1999; 5:662–8. [PubMed: 10371505]

50. Zhang H, et al. Beta 1-integrin protects hepatoma cells from chemotherapy induced apoptosis via a mitogen-activated protein kinase dependent pathway. Cancer. 2002; 95:896–906. [PubMed: 12209735]

51. Argyris EG, et al. The interferon-induced expression of APOBEC3G in human blood-brain barrier exerts a potent intrinsic immunity to block HIV-1 entry to central nervous system. Virology. 2007; 367:440–51. [PubMed: 17631933]

52. Shain KH, Dalton WS. Cell adhesion is a key determinant in de novo multidrug resistance (MDR): new targets for the prevention of acquired MDR. Mol Cancer Ther. 2001; 1:69–78. [PubMed: 12467240]

53. Damiano JS. Integrins as novel drug targets for overcoming innate drug resistance. Curr Cancer Drug Targets. 2002; 2:37–43. [PubMed: 12188919]

54. Hazlehurst LA, Dalton WS. Mechanisms associated with cell adhesion mediated drug resistance (CAM-DR) in hematopoietic malignancies. Cancer Metastasis Rev. 2001; 20:43–50. [PubMed: 11831646]

55. Hehlgans S, Haase M, Cordes N. Signalling via integrins: implications for cell survival and anticancer strategies. Biochim Biophys Acta. 2007; 1775:163–80. [PubMed: 17084981]

56. Cho S, et al. Generation, characterization and preclinical studies of a human anti-L1CAM monoclonal antibody that cross-reacts with rodent L1CAM. MAbs. 2016; 8:414–25. [PubMed: 26785809]

57. Min JK, et al. L1 cell adhesion molecule is a novel therapeutic target in intrahepatic cholangiocarcinoma. Clin Cancer Res. 2010; 16:3571–80. [PubMed: 20501614]

58. Schäfer H, et al. Combined treatment of L1CAM antibodies and cytostatic drugs improve the therapeutic response of pancreatic and ovarian carcinoma. Cancer Lett. 2012; 319:66–82. [PubMed: 22210381]

59. Golubovskaya VM. Targeting FAK in human cancer: from finding to first clinical trials. Front Biosci (Landmark Ed). 2014; 19:687–706. [PubMed: 24389213]

60. Schultze A, Fiedler W. Therapeutic potential and limitations of new FAK inhibitors in the treatment of cancer. Expert Opin Investig Drugs. 2010; 19:777–88.

61. Rezaee M, Sanche L, Hunting DJ. Cisplatin enhances the formation of DNA single- and double-strand breaks by hydrated electrons and hydroxyl radicals. Radiat Res. 2013; 179:323–31. [PubMed: 23368416]

62. Zellweger R, et al. Rad51-mediated replication fork reversal is a global response to genotoxic treatments in human cells. J Cell Biol. 2015; 208:563–79. [PubMed: 25733714]

63. Roberts SA, Gordenin DA. Hypermutation in human cancer genomes: footprints and mechanisms. Nat Rev Cancer. 2014; 14:786–800. [PubMed: 25568919]

64. Chan K, et al. An APOBEC3A hypermutation signature is distinguishable from the signature of background mutagenesis by APOBEC3B in human cancers. Nat Genet. 2015; 47:1067–72. [PubMed: 26258849]

65. Lipinski KA, et al. Cancer Evolution and the Limits of Predictability in Precision Cancer Medicine. Trends Cancer. 2016; 2:49–63. [PubMed: 26949746]

66. Beltran H, et al. Whole-Exome Sequencing of Metastatic Cancer and Biomarkers of Treatment Response. JAMA Oncol. 2015; 1:466–74. [PubMed: 26181256]

67. Demichelis F, et al. SNP panel identification assay (SPIA): a genetic-based assay for the identification of cell lines. Nucleic Acids Res. 2008; 36:2446–56. [PubMed: 18304946]

68. Beltran H, et al. Divergent clonal evolution of castration-resistant neuroendocrine prostate cancer. Nat Med. 2016; 22:298–305. [PubMed: 26855148]

69. Futreal PA, et al. A census of human cancer genes. Nat Rev Cancer. 2004; 4:177–83. [PubMed: 14993899]

70. Rubio-Perez C, et al. In silico prescription of anticancer drugs to cohorts of 28 tumor types reveals targeting opportunities. Cancer Cell. 2015; 27:382–96. [PubMed: 25759023]

71. Cibulskis K, et al. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. Nat Biotechnol. 2013; 31:213–9. [PubMed: 23396013]

72. Jiang Y, Soong TD, Wang L, Melnick AM, Elemento O. Genome-wide detection of genes targeted by non-Ig somatic hypermutation in lymphoma. PLoS One. 2012; 7:e40332. [PubMed: 22808135]

73. Romanel A, Lago S, Prandi D, Sboner A, Demichelis F. ASEQ: fast allele-specific studies from next-generation sequencing data. BMC Med Genomics. 2015; 8:9. [PubMed: 25889339]

74. Ramos AH, et al. Oncotator: cancer variant annotation tool. Hum Mutat. 2015; 36:E2423–9. [PubMed: 25703262]

75. Eilbeck K, et al. The Sequence Ontology: a tool for the unification of genome annotations. Genome Biol. 2005; 6:R44. [PubMed: 15892872]
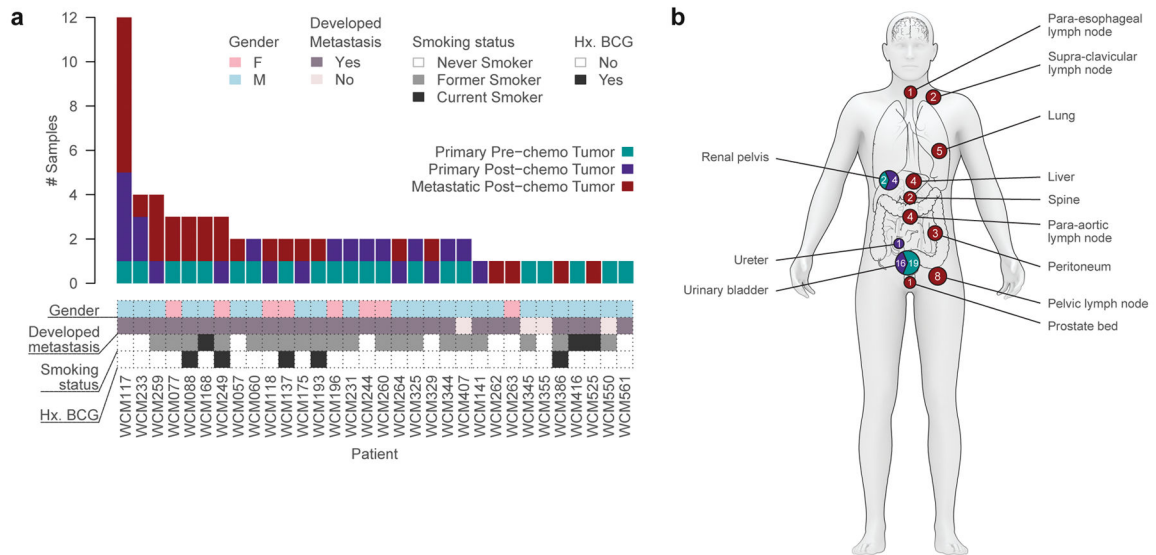
Figure 1. Clinical characteristics of study cohort
(a) Bar graph illustrating number of tumor samples sequenced from each study subject.
Treatment status of each sample is color-coded within each bar (see inset). Gender, smoking
status development of metastases and history of intravesical Bacillus Calmette-Guerin
(BCG) for pre-existing non-muscle invasive bladder cancer are represented on the bottom
(b) Schematic illustrating anatomical sites of primary and metastatic tumor samples.

Figure 2. Clonal mutational heterogeneity in chemotherapy-treated UC
(a) Percentage of shared and unique SNVs in matched chemotherapy-naive and chemotherapy-treated UC tumors. **(b)** Discordance in the mutational status in selected driver genes between chemotherapy-naive and chemotherapy-treated UC tumors. Each column represents a paired set of pre- and post-chemotherapy tumors obtained from an individual subject. Individuals with more than one paired set of tumors are highlighted with matching color designations. Primary-primary tumors are represented on the left (vertical hatching) and primary-metastatic tumors are represented on the right (slanted hatching). Note: *ERCC2* missense mutations were identified in two tumor samples, WCM259_3 (p.Q755H) and WCM141_1 (p.T484M), matched pre- and post-chemotherapy samples were not available from either patient.
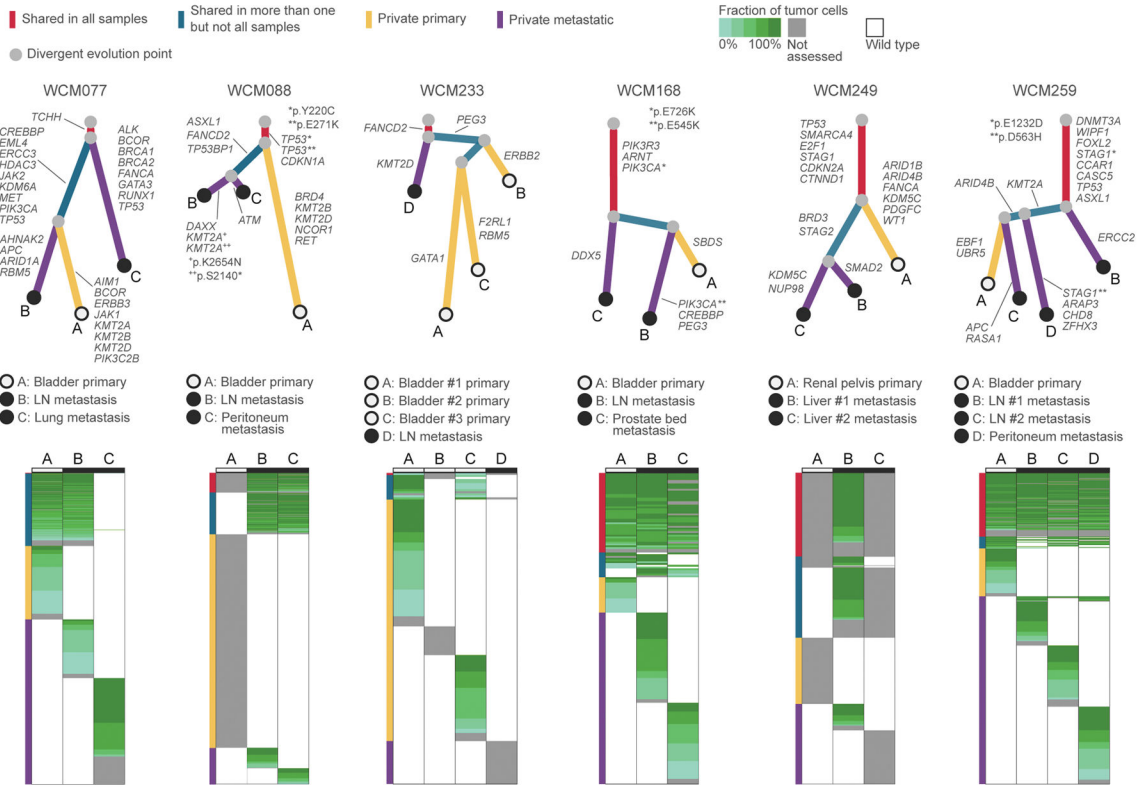
Figure 3. Early branching evolution in UC

Phylogenetic trees (top), shared and private clonally-adjusted mutations (bottom) from 6 patients with three or more tumor samples per patient. All cases showed early metastatic spread occurring as branching evolution occurred early in the natural history of the disease. All primary tumors temporally preceded the development of metastasis but were assigned "branch" values based on genomic distance suggesting early branching evolution and metastatic spread.
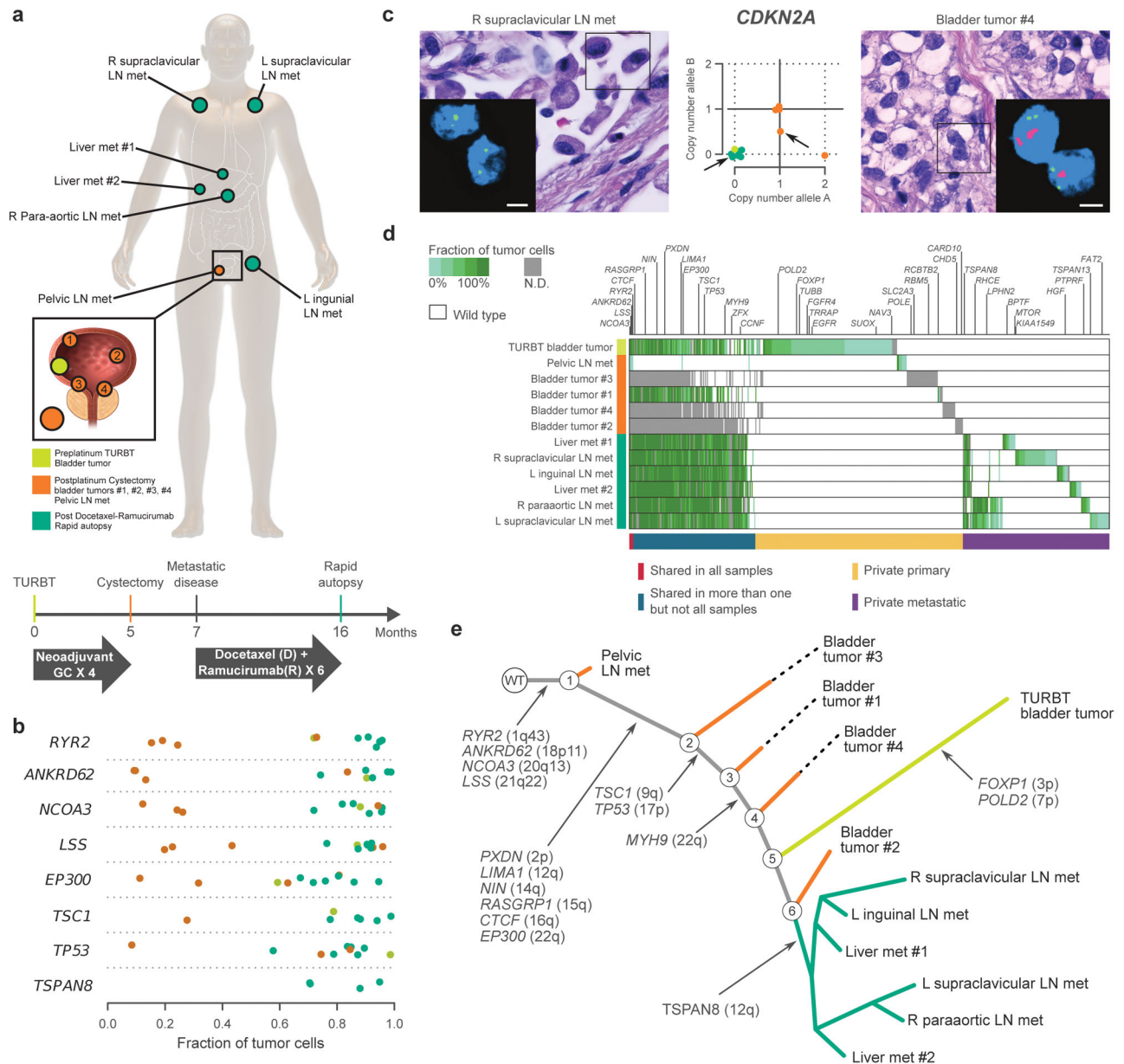
Figure 4. Reconstructing the spatio-temporal evolution of UC over time and through different treatments

Analysis of 12 tumor samples collected during disease progression and at the time of death of patient WCM117 **(a)** (top) Circles represent sites of sequenced tumors. (bottom) Timeline and clinical course (vertical lines) in the natural history of the disease. **(b)** CLONET-adjusted variant allele frequencies of selected mutations. Each dot represents the fraction of cells harboring the corresponding mutation. **(c)** Fluorescence *in-situ* hybridization for *CDKN2A* (red) and reference (green) probes in primary cystectomy tumor #4 (right) and right supraclavicular lymph node collected at time of autopsy (left) (scale bar 5 μm). Middle panel: allele-specific copy number. Axes correspond to different alleles. Dots represent *CDKN2A* allele-specific copy number. **(d)** CLONET-adjusted shared and private mutations.

Fractions of tumor cells harboring each mutation represented by shades of green (scale upper left corner of panel). Gray: clonality information was not available. **(e)** Reconstruction of evolutionary tree. WT: taken last ancestor with wild-type genome that acquires a series of mutations during oncogenesis. The length of the branches represents the distance between two tumors based on the number of shared mutations. Samples are color-coded consistently across panels: Yellow: primary pre-chemotherapy TURBT. Orange: 4 different areas of post-gemcitabine cisplatin bladder tumors and pelvic lymph node metastatic lesion removed through cystectomy and lymph node dissection. Green: post-docetaxel-ramucirumab metastatic lesions removed during autopsy.
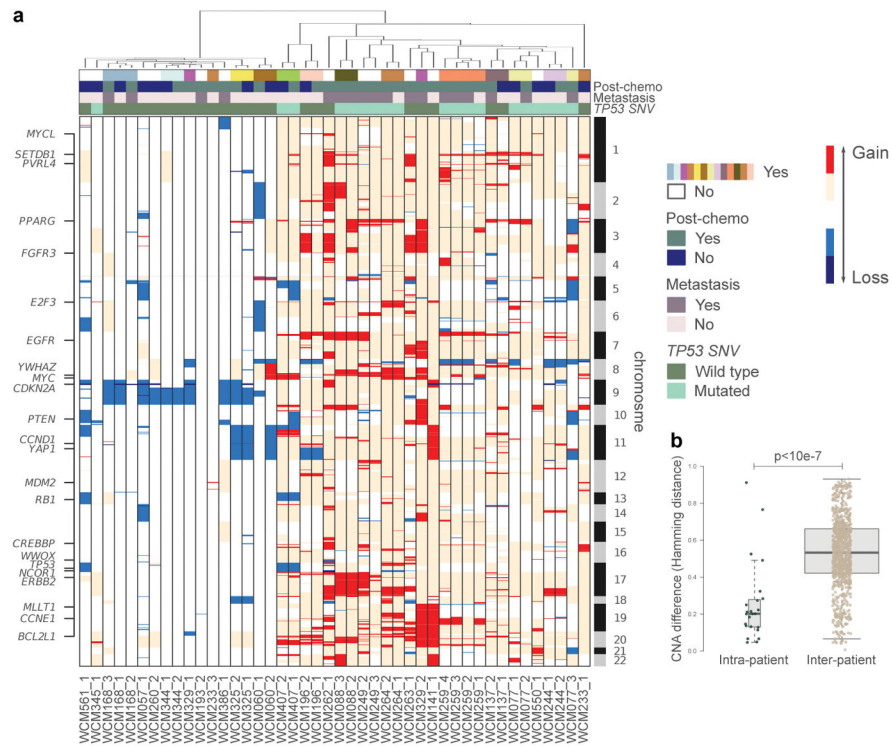
Figure 5. Hierarchical clusters of 44 UC tumor samples by copy-number alterations
(a) Copy number gains are represented in red and copy number losses are represented in blue. Each column represents one tumor sample. Clinical annotations are represented on top. Matched samples from the same patient are represented in the same color in the "matched samples" annotation track. Selected genes harboring frequent copy number alterations are listed on the left. **(b)** Dots represent pairs of samples in Fig. 5a from the same patient (left) and from different patients (right). Left boxplot: median 0.2, IQR (0.13, 0.28), whiskers (0.05, 0.49). Right boxplot: median 0.53, IQR (0.42, 0.66), whiskers (0.07, 0.93).
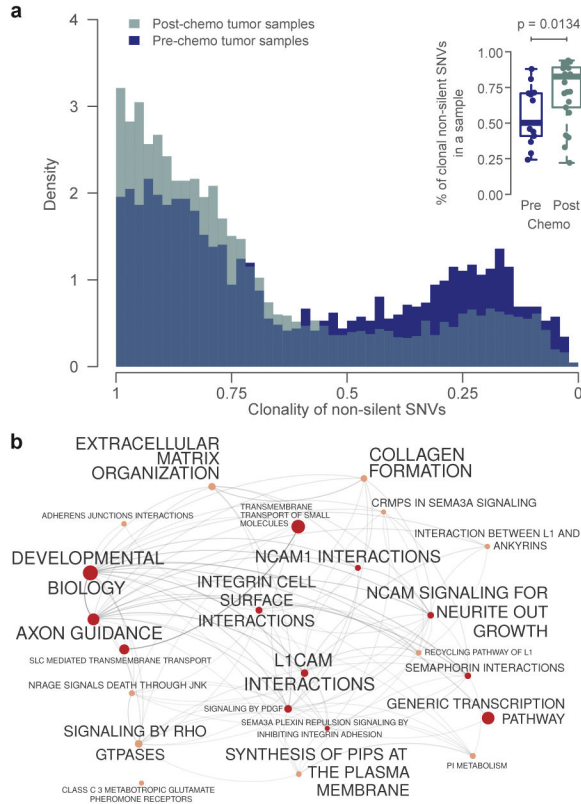
Figure 6. Clonal enrichment of mutations in chemotherapy-treated UC
(a) Density plot representing clonality of non-silent single nucleotide variants (SNVs) on the X-axis and the density distribution on the Y-axis. Pre-chemotherapy tumor samples are represented in blue color and post-chemotherapy tumor samples in green. Box plots (right) represent the percentage of clonal SNVs in pre-chemotherapy tumor and post-chemotherapy tumors demonstrating a significant increase in clonality in post-chemotherapy UC tumors. Left boxplot: median 0.5, IQR (0.41, 0.71), whiskers (0.24, 0.88). Right boxplot: median 0.83, IQR (0.61, 0.89), whiskers (0.22, 0.94). **(b)** Results of Gene set Enrichment Analysis (GSEA) of mutations in chemotherapy-treated UC. Each node represents one REACTOME pathway. Node size represents number of genes in each pathway. Enriched pathways post-chemotherapy samples are represented by nodes in red color.
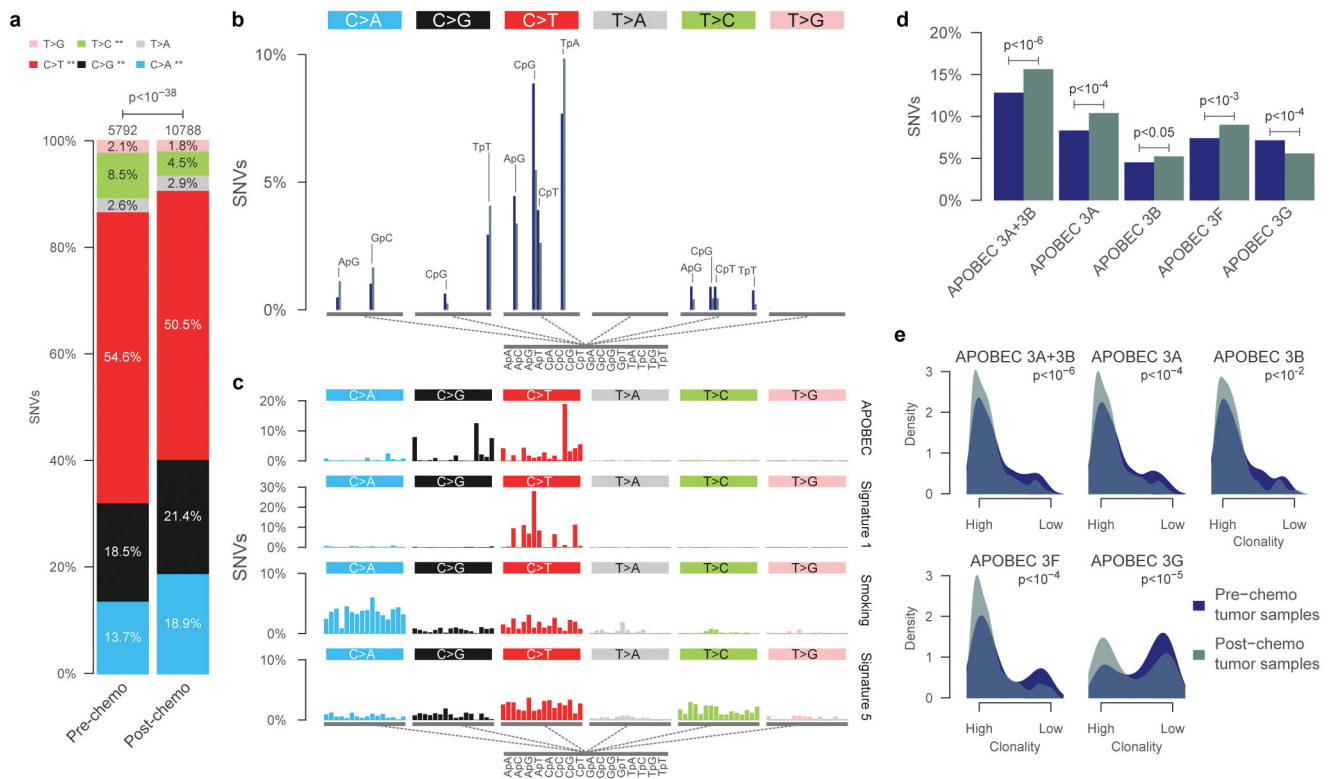
Figure 7. Mutagenesis in advanced UC is shaped by chemotherapy and APOBECs
(a) Composite bar graphs representing the distribution of all possible nucleotide substitutions in non-silent SNVs in sequenced pre-chemotherapy (left) and post-chemotherapy (right) UC tumors. **(b)** Nucleotide motif contexts for each category of single nucleotide substitutions. Blue: pre-chemotherapy tumors, green: post-chemotherapy tumors. **(c)** Four mutational signatures identified in UC. **(d)** Significant enrichment of APOBEC signatures in post-chemotherapy tumors. **(e)** Significant increases in the clonality of APOBEC-induced mutations in post-chemotherapy UC. Number of SNVs, exact p-values, and total number of samples are reported in Supplementary Note.