**Title:** Somatic ERCC2 mutations correlate with cisplatin sensitivity in muscle-invasive urothelial carcinoma

**Authors:**
Eliezer M. Van Allen1,2*, Kent W. Mouw3,4*, Philip Kim5, Gopa Iyer7,9, Nikhil Wagle1,2, Hikmat Al-Ahmadie6,7, Cong Zhu2, Irina Ostrovnaya8, Gregory V. Kryukov2, Kevin W O’Connor2, John Sfakianos9, Ilana Garcia-Grossman9, Jaegil Kim2, Elizabeth A. Guancial10, Richard Bambury6, Samira Bahl2, Namrata Gupta2, Deborah Farlow2, Angela Qu1, Sabina Signoretti11, Justine A. Barletta11, Victor Reuter6,7, Jesse Boehm2, Michael Lawrence2, Gad Getz2,12, Philip Kantoff1, Bernard H. Bochner5,7, Toni K. Choueiri1, Dean F. Bajorin7,8, David B. Solit7,9,12, Stacey Gabriel1, Alan D’Andrea3,4, Levi A. Garraway1,2**, and Jonathan E. Rosenberg7,9**

**Affiliations:**
1) Department of Medical Oncology, Dana-Farber Cancer Institute, Harvard Medical School, 450 Brookline Avenue, Boston, MA 02115
2) Broad Institute of MIT and Harvard, 7 Cambridge Center, Cambridge, MA 02142
3) Department of Radiation Oncology, Dana-Farber Cancer Institute, Harvard Medical School, 450 Brookline Avenue, Boston MA 02115
4) Harvard Radiation Oncology Program, Boston, MA 02115, USA
5) Urology Service, Department of Surgery, Memorial Sloan-Kettering Cancer Center, 1275 York Avenue, New York, NY 10065
6) Department of Pathology, Memorial Sloan-Kettering Cancer Center, 1275 York Avenue, New York, NY 10065
7) Weill Cornell Medical College, Cornell University, New York, NY 10065
8) Department of Epidemiology and Biostatistics, Memorial Sloan-Kettering Cancer Center, 1275 York Avenue, New York, NY 10065
9) Genitourinary Oncology Service, Department of Medicine, Memorial Sloan-Kettering Cancer Center, 1275 York Avenue, New York, NY 10065
10) Division of Hematology/Oncology, Department of Medicine, University of Rochester Medical Center School of Medicine and Dentistry, 601 Elmwood Ave, Box 704, Rochester, NY 14642
11) Department of Pathology, Brigham and Women’s Hospital, Harvard Medical School, 75 Francis Street, Boston MA
12) Massachusetts General Hospital Cancer Center and Department of Pathology, Boston, MA 02114
13) Human Oncology and Pathogenesis Program, Memorial Sloan-Kettering Cancer Center, 1275 York Avenue, New York, NY 10065

**Running Title**
ERCC2 and cisplatin sensitivity in urothelial carcinoma

**Keywords**
ERCC2, cisplatin sensitivity, urothelial carcinoma, exceptional responders, nucleotide excision repair defect

**Corresponding Authors:**
Jonathan E. Rosenberg
Genitourinary Oncology Service
Memorial Sloan-Kettering Cancer Center
1275 York Avenue
New York, NY 10065
Phone: 646-422-4461
Fax: 646-227-2417
rosenbj1@mskcc.org

Levi A. Garraway
Department of Medical Oncology
Dana-Farber Cancer Institute
450 Brookline Avenue, D1542
Boston, MA 02115, USA
Phone: 617-632-6689
Fax: 617-582-7880
levi_garraway@dfci.harvard.edu

*These authors contributed equally to this manuscript
**These authors contributed equally to this manuscript

Word Count: 4,612

Number of Figure and Tables: 7


Conflict of Interest Statement: Dr. Rosenberg is a consultant for Boehringer Ingelheim, Bristol Myers Squib, Oncogenex, Onyx, Johnson and Johnson, and Dendreon. Dr. Garraway and Dr. Wagle are equity holders in and consultants to Foundation Medicine. Dr. Garraway is a consultant to Novartis, Millenium/Takeda, and Boehringer Ingelheim, and a recipient of a sponsored research grant from Novartis. Drs. Rosenberg, Garraway, Van Allen, Mouw, Wagle, and D’Andrea have a patent pending for the relationship between somatic ERCC2 mutations and cisplatin response.
ABSTRACT
Cisplatin-based chemotherapy is the standard of care for patients with muscle invasive urothelial carcinoma. Pathologic downstaging to pT0/pTis after neoadjuvant cisplatin-based chemotherapy is associated with improved survival, although molecular determinants of cisplatin response are incompletely understood. We performed whole exome sequencing on pre-treatment tumor and germline DNA from 50 patients with muscle invasive urothelial carcinoma who received neoadjuvant cisplatin-based chemotherapy followed by cystectomy (25 pT0/pTis “responders”, 25 pT2+ “non-responders”) to identify somatic mutations that occurred preferentially in responders. ERCC2, a nucleotide excision repair gene, was the only significantly mutated gene enriched in the cisplatin responders compared with non-responders (q < 0.01). Expression of representative ERCC2 mutations in an ERCC2-deficient cell line failed to rescue cisplatin and UV sensitivity compared to wild-type ERCC2. Lack of normal ERCC2 function may contribute to cisplatin sensitivity in urothelial cancer and somatic ERCC2 mutation status may inform cisplatin-containing regimen usage in muscle invasive urothelial carcinoma.

STATEMENT OF SIGNIFICANCE
Somatic ERCC2 mutations correlate with complete response to cisplatin-based chemosensitivity in muscle-invasive urothelial carcinoma and clinically identified mutations lead to cisplatin-sensitivity in vitro. Nucleotide excision repair pathway defects may drive exceptional response to conventional chemotherapy.
INTRODUCTION

Platinum-based chemotherapy has been the standard of care for patients with muscle invasive and metastatic urothelial carcinoma for over 20 years(1-3), and neoadjuvant cisplatin-based chemotherapy leads to a 14-25% relative risk reduction for death from muscle invasive urothelial carcinoma (cT2-T4aN0M0)(3-5). Pathologic downstaging to complete response (pT0) or carcinoma in situ (pTis) at cystectomy occurs in 26-38% of patients treated with neoadjuvant chemotherapy compared to 12.3-15% for patients undergoing cystectomy alone(3, 4, 6), and the 5-year survival for pT0/pTis patients is 85% after neoadjuvant chemotherapy(3) compared to 43% for patients with persistent muscle invasive disease (≥ pT2)(7). Therefore, the benefit of neoadjuvant chemotherapy appears to be most dramatic in patients who are found to have pathological complete responses at the time of surgical resection. However, the inability to predict which patients will derive clinical benefit has limited the use of this toxic approach in the urologic community(8-10).

Cisplatin causes accumulation of DNA crosslinks, which interfere with DNA replication and gene transcription, and eventually promote cell death. Repair of cisplatin-induced DNA damage occurs primarily through DNA repair pathways such as nucleotide excision repair (NER) (11) and homologous recombination (which includes BRCA1 and BRCA2).(12). The NER pathway involves multiple genes, including ERCC1-5, CDK7, DDB1-2, XPA, and XPC.(13) Germline alterations in NER genes result in multiple recessive inherited disorders, including xeroderma pigmentosum (XP)(13). Due to a deficiency in NER, XP patients have significantly increased risk of developing skin cancers and other malignancies(14).
Germline or expression-based changes in several NER genes have been suggested to modulate clinical response to cisplatin-based chemotherapy(15, 16). However, prospective studies have not confirmed these observations(17, 18). ERCC2 (an NER helicase) loss-of-function correlates with cisplatin sensitivity in preclinical models(19), while ERCC2 overexpression leads to cisplatin resistance(20). These data suggest that tumors with loss of NER function may exhibit increased cisplatin sensitivity, and recent data has identified somatic ERCC2 mutations in 7-12% of urothelial carcinomas(21, 22). We hypothesized that somatic mutations in the NER pathway may correlate with response to cisplatin-based neoadjuvant chemotherapy in urothelial carcinoma patients. To test this, and more generally to define genomic correlates of chemotherapy response, we performed whole exome sequencing of urothelial carcinoma tumors from patients with extreme responses to neoadjuvant cisplatin-based combination chemotherapy.

RESULTS

Somatic genetic alterations in muscle invasive urothelial carcinoma

We sequenced pre-treatment tumor and germline DNA from 50 patients treated with neoadjuvant cisplatin-based chemotherapy for muscle invasive urothelial carcinoma; 25 “responders” had no residual invasive disease (pT0/pTis) on pathologic examination following cystectomy, and 25 “non-responders” had residual muscle-invasive (≥ pT2) disease (Fig. 1A, Methods). Although multiple chemotherapeutic regimens were utilized, the only common agent among all patients was cisplatin (Table 1, Supplementary Table 1). No significant differences in clinical characteristics were identified between responders and non-responders at baseline (P > 0.05; Mann-Whitney).
The mean target coverage from whole exome sequencing (WES) was 121X for tumors and 130X for paired germline samples (Supplementary Table 1). The median mutation rate was 9.7 mutations per megabase (mutations/Mb) for responders and 4.4 mutations/Mb for non-responders ($P = 0.0003$; Mann-Whitney test) (Fig. 1B, S1A-B, Supplementary Table 1), raising the possibility of reduced DNA repair fidelity among cisplatin responders. All observed somatic mutations and short insertion/deletions are reported in Supplementary Table 2.

A statistical assessment (23) of the base mutations and short insertion/deletions across both responders and non-responders nominated four significantly altered genes previously implicated in urothelial carcinoma (21, 22, 24): TP53, RB1, KDM6A, and ARID1A (Fig. 1B, Supplementary Table 3-4). In addition, nine non-synonymous somatic mutations were observed in ERCC2 (Fig. 1B, S2; Supplementary Table 2). Although ERCC2 did not reach cohort-wide statistical significance, its known role in DNA repair and report of being recurrently mutated in bladder cancer (21, 22) raised the possibility that ERCC2 mutations might associate with cisplatin response.

Somatic ERCC2 mutations in cisplatin-based chemotherapy responders

We performed an enrichment analysis to identify genes that were selectively mutated in the responders compared to non-responders (Online Methods). Among 3,277 genes with at least one possibly damaging somatic alteration (Online Methods), ERCC2 was the only gene significantly enriched in the responder cohort (Fig. 2A, Supplementary Tables 5-6). Indeed, all ERCC2 non-synonymous somatic mutations occurred in the cisplatin sensitive tumors ($P < 0.001$; Fisher’s exact test). ERCC2 remained significantly enriched in responders following false discovery analysis performed on genes in which the mutation frequency afforded adequate power ($q = 0.007$; Benjamini-Hochberg) (Fig.
Moreover, the enrichment for ERCC2 mutations in the responder group was significant when adjusted for differences in overall mutation rate between responders and non-responders \((P = 0.04; \text{binomial test})\). Towards this end, the median background mutation rate for ERCC2 mutant tumors (15.5 mutations per megabase) was significantly elevated compared to ERCC2 wild-type tumors (5.1 mutations per megabase) \((P = 0.01; \text{Mann-Whitney test})\) (Supplementary Fig. 1B), consistent with a possible DNA repair defect and prior reports(22).

The somatic ERCC2 mutation frequency in the responder cohort was also compared to two unselected bladder cancer populations: 130 cases from the Cancer Genome Atlas (TCGA)(21) and 99 cases from a Chinese patient cohort(22) (Fig. 2C). When compared to these unselected populations, ERCC2 mutations were significantly enriched in the cisplatin responder cohort (36% of cases; \(P < 0.001; \text{binomial test})\) (Fig. 2C). ERCC2 mutation status does not appear to be prognostic, as it had no impact on overall survival in the TCGA cohort \((P = 0.77; \text{Log-Rank})\) (Supplementary Fig. 3). To determine the relative abundance of somatic ERCC2 mutations in other tumor types, TCGA data from 19 tumor types \((n = 4,429)\) was queried(25). Somatic ERCC2 mutations were observed at low frequencies (< 4%) in 11 other tumor types (Fig. 3A-B).

All identified somatic ERCC2 mutations occurred at highly conserved amino acid positions within or immediately adjacent to the helicase domains (Fig. 3A, C, S4). Similarly, germline ERCC2 mutations in patients with XP (complementation group D) and XP with combined Cockayne syndrome (XP/CS), two disorders characterized by NER function, cluster within helicase domains (Fig. 3C). Conversely, mutations causing trichothiodystrophy, a disease resulting from alteration of ERCC2’s role in transcription, are distributed throughout the protein(26).
ERCC2 mutations confer increased cisplatin sensitivity in vitro

These observations raised the possibility that the identified mutations disrupt ERCC2 function and interfere with NER. To test this hypothesis, the first five of the identified ERCC2 mutants were stably expressed in an immortalized ERCC2-deficient cell line derived from an XP patient, and the cisplatin sensitivity profile of each cell line was measured (Online Methods, Fig. 4A). Expression of wild-type (WT) ERCC2 rescued cisplatin sensitivity of the ERCC2-deficient cell line, whereas none of the ERCC2 mutants rescued cisplatin sensitivity (Fig. 4B). The IC₅₀ for the ERCC2WT-complemented cell line was significantly higher than the ERCC2-deficient parent cell line (P < 0.0001; ANOVA), whereas the IC₅₀ for each ERCC2-mutant complemented cell line was not significantly different than the parent ERCC2-deficient parent cell line (Fig 4C). Similarly, the IC₅₀ for the ERCC2WT cell line was significantly higher than the ERCC2-deficient and mutant cell lines (P < 0.0001; ANOVA).

The NER pathway repairs DNA lesions other than cisplatin adducts, so we also tested the effect of the identified ERCC2 mutations on NER-mediated repair of UV-induced DNA damage. WT and mutant ERCC2 complemented cell lines were exposed to increasing doses of UV irradiation, and clonogenic survival was measured (Fig. 5A-C). Whereas the ERCC2WT-complemented cell line rescued UV sensitivity, the UV sensitivities of the ERCC2-mutant complemented cell lines were not significantly different than that of the ERCC2-deficient parent cell line. Taken together, these experiments suggest that the observed ERCC2 mutations result in loss of normal NER capacity.

Since the overall mutation rate was higher in ERCC2-mutated tumors, we hypothesized that ERCC2 mutations may be broadly contributing to genomic instability. Thus, we
measured rates of chromosomal aberrations in WT and mutant ERCC2 complemented cell lines before and after cisplatin treatment. In the absence of cisplatin, background rates of chromosomal aberrations were slightly lower in the ERCC2\textsuperscript{WT}-complemented cell line than in the ERCC2-deficient or mutant complemented cell lines, but this difference was not statistically significant (Fig 6A-D). However, following cisplatin exposure, significantly fewer chromosomal aberrations were observed in the ERCC2\textsuperscript{WT}-complemented cell line compared to the ERCC2-deficient parent cell line ($P = 0.03$; ANOVA), whereas expression of the ERCC2 mutants resulted in no rescue of chromosomal stability ($P > 0.5$) (Fig. 6D). These data suggest that the responder-associated ERCC2 mutations may contribute to overall genomic instability in these tumors.

Other DNA repair gene alterations

Finally, we sought to determine if other DNA repair genes might undergo recurrent mutations in cisplatin-sensitive tumors. No significantly recurrent mutations were observed in other NER or homologous repair genes in responders compared to non-responders. However, in two responder tumors that did not have ERCC2 mutations, somatic nonsense (truncating) mutations were detected in the homologous recombination DNA repair genes BRCA1 and BRCA2 (Supplementary Table S2). There were no nonsynonymous BRCA1 or BRCA2 mutations in the non-responders (Supplementary Table S2). These results are consistent with the known sensitivity of BRCA-mutant tumors (e.g. BRCA1/2-mutant breast or ovarian cancer) to platinum-containing regimens(27). While somatic ERCC2 mutations were the only DNA repair gene mutations enriched in the responders, singleton missense mutations of uncertain significance were observed in DNA damage response genes in both responders and non-responders, and are of unknown functional relevance.
DISCUSSION

In bladder cancer, the clinical benefit of neoadjuvant chemotherapy is most apparent when pathological downstaging to pT0 or pTis is achieved at surgical resection following cisplatin-based chemotherapy. However, the lack of a predictive biomarker for clinical benefit from neoadjuvant cisplatin-based chemotherapy has limited the use of this approach in the clinical community due to toxicity concerns. Using an extreme phenotype analysis, we have identified an association between somatic ERCC2 mutations and pathologic complete response to neoadjuvant cisplatin-based chemotherapy in muscle invasive urothelial carcinoma. While ERCC2 mutations occur in approximately 12% of unselected cases, 36% of cisplatin-based chemotherapy responders in our cohort harbored somatic ERCC2 non-synonymous mutations. Moreover, all ERCC2 mutant tumors responded to neoadjuvant chemotherapy.

The NER pathway is a highly conserved DNA repair system that identifies and repairs bulky DNA adducts arising from genotoxic agents such as cisplatin. The NER helicase ERCC2 unwinds duplex DNA near the damage site through the coordinated action of two conserved helicase domains. All ERCC2 mutations identified in this study occurred at conserved positions within or adjacent to these helicase domains, and the identified mutants all failed to complement cisplatin or UV sensitivity of an ERCC2-deficient cell line. Together, these data suggest that the mutations result in loss of normal ERCC2 function, leading to increased tumor cell sensitivity to DNA-damaging agents such as cisplatin.

Interestingly, in seven (78%) of the ERCC2-mutant cases, the ERCC2 mutation allelic fraction was < 0.5 (Supplementary Table S2), suggesting that WT ERCC2 remains present at one allele. Therefore, the cisplatin sensitivity phenotype may result from a haploinsufficient or dominant negative effect of a heterozygous ERCC2 mutation, rather
as a result of biallelic inactivating mutations (as in the traditional “two-hit” tumor suppressor model). The driving force for heterozygous mutation of ERCC2 is unknown; however, the prevalence of ERCC2 mutations in this study and other cohorts (such as the TCGA) suggests that loss of normal ERCC2 function may provide a selective advantage for tumors. Partial loss of DNA repair fidelity could aid tumor growth by decreasing repair-associated delays in cell cycle progression. In addition, decreased NER capacity may result in higher rates of error-prone repair or large-scale genomic changes that further drive tumor growth.

Despite providing a potential growth advantage, mutation of one ERCC2 allele may render tumor cells susceptible to DNA damaging agent such as cisplatin if inadequate levels of WT ERCC2 are present to support NER (i.e. haploinsufficiency). Alternatively, the mutated version of ERCC2 may bind but not efficiently repair damaged DNA, thereby preventing repair by an alternative DNA repair pathway and leading to a dominant-negative phenotype, as has been described for mutants of the yeast ERCC2 homolog, Rad3(28). Further studies are necessary to explore the effects of ERCC2 loss on tumor growth, and the mechanism by which the identified ERCC2 mutations confer changes in tumor NER capacity.

One possible explanation for the findings observed in this study is that somatic ERCC2 mutation is associated with good prognosis small tumors. However, this is not supported by the survival data from the TCGA bladder cancer cohort, which excluded patients who received neoadjuvant chemotherapy. No difference in overall survival was observed based on somatic ERCC2 mutation status (Fig. S3). In the extreme response cohort reported here, patients were generally noted to have obvious disease left behind after initial transurethral resection based on operative reports and imaging, making the findings unlikely related to complete transurethral resection.
While these data should not yet be used to justify avoiding cisplatin-based treatment in ERCC2 WT patients, our findings raise the possibility that somatic ERCC2 mutation status may provide a genetic means to select patients most likely to benefit from cisplatin-based chemotherapy, while directing other patients towards alternative therapeutic approaches. Of note, our study focused specifically on somatic mutations that are exclusively in the tumor, and not germline single nucleotide polymorphisms in ERCC2 or other genes. Thus, this approach is distinct from genome-wide association studies that have examined germline ERCC1 or ERCC2 polymorphisms and their mixed impact on cisplatin sensitivity (29). Broadly, these findings will require independent clinical validation in prospective trials to establish the clinical predictive power of somatic ERCC2 mutation status for cisplatin response.

It is possible that some nonresponding urothelial tumors will harbor somatic ERCC2 mutations in larger cohorts; if observed, examination of the post-chemotherapy resistant tumor would be critical to understand whether tumor heterogeneity played a role in resistance. The patients analyzed in this extreme phenotype analysis were treated with combination cisplatin-based chemotherapy regimens containing non-cisplatin agents; however, cisplatin was the only agent common to all regimens. Since half of patients with bladder cancer are not candidates for cisplatin-based chemotherapy due to pre-existing comorbidities, less toxic carboplatin-based neoadjuvant therapies may warrant study for non-cisplatin eligible patients with ERCC2-mutant tumors.

To date, exceptional response genomic studies have informed genomic mechanisms of response to targeted therapies, such as response to everolimus in multiple disease contexts (30, 31). However, published studies have been limited to individual case reports due to the rarity of such events with targeted therapies. This study represents a new approach for studying extraordinary responses to commonly used cytotoxic
chemotherapies using a case-control design, which may be applied to other therapeutic settings in which a significant minority of patients achieve exceptional response (e.g. neoadjuvant chemotherapies in other clinical settings). Towards that end, a majority of responder cases in our cohort had no recurrent genomic determinant of cisplatin response. It is possible that alterations in DNA repair genes not readily detectable with WES (e.g. epigenetic, expression-based) may mediate cisplatin sensitivity in these cases.

In conclusion, this work provides new insights into the relationship between somatic genetic alterations and clinical response to cisplatin-based chemotherapy in urothelial carcinoma. If further validated, these results may inform therapeutic decision-making, novel therapeutic development, and clinical trial designs for urothelial carcinoma and possibly other ERCC2-mutated tumors. Finally, these results show that somatic genomic alterations may reveal the mechanistic underpinnings of anti-tumor response to conventional cytotoxic chemotherapy.

METHODS

Patients and Samples

Patients with muscle invasive or locally advanced urothelial carcinoma and extreme responses to chemotherapy (defined as no residual invasive carcinoma at cystectomy or presence of persistent muscle invasive or extravesical disease at cystectomy), available pre-chemotherapy tumor tissue, and enrolled on Institutional Review Board (IRB) approved tissue acquisition and DNA sequencing protocols were identified (Dana-Farber protocols 02-021 and 11-334; Memorial Sloan-Kettering Cancer Center protocols 89-076 and 09-025). All patients provided written informed consent for genomic testing utilized for this study. Specimens were evaluated by genitourinary pathologists (JAB/SS - DFCI...
cohort, and HA - MSKCC cohort) to identify tumor-bearing areas for DNA extraction. The minimum percentage of neoplastic cellularity for regions of tumor tissue was 60%. Study specimens of frozen or formalin-fixed, paraffin-embedded (FFPE) tissue sections were identified at the Dana-Farber Cancer Institute and Memorial Sloan-Kettering Cancer Center. Germline DNA was extracted either from peripheral blood mononuclear cells or histologically normal non-urothelial tissue. Information about the source of germline DNA is available in Supplementary Table 1.

**Whole Exome Sequencing and Statistical Analysis**

DNA extraction and exome sequencing: Slides were cut from FFPE or frozen tissue blocks and examined by a board-certified pathologist to select high-density cancer foci and ensure high purity of cancer DNA. Biopsy cores were taken from the corresponding tissue block for DNA extraction. DNA was extracted using Qiagen's QIAamp DNA FFPE Tissue Kit Quantitation Reagent (Invitrogen). DNA was stored at -20 degrees Celsius. Whole exome capture libraries were constructed from 100ng of DNA from tumor and normal tissue after sample shearing, end repair, and phosphorylation and ligation to barcoded sequencing adaptors. Ligated DNA was size selected for lengths between 200-350 bp and subjected to exonic hybrid capture using SureSelect v2 Exome bait (Agilent). The sample was multiplexed and sequenced using Illumina HiSeq technology for a mean target exome coverage of 121X for the tumors and 130X for germline samples. Four cases did not complete the exome sequencing process due to sequencing process failure. All BAM files generated for this study will be deposited in dbGap (submission in progress).

Sequence data processing: Exome sequence data processing and analysis were performed using pipelines at the Broad Institute. A BAM file aligned to the hg19 human
genome build was produced using Illumina sequencing reads for the tumor and normal sample and the Picard pipeline. BAM files were uploaded into the Firehose infrastructure (32), which managed intermediate analysis files executed by analysis pipelines. All BAM files will be uploaded to dbGaP (phs000771.v1.p1).

Sequencing quality control: Sequencing data was incorporated into quality control modules in Firehose to compare the tumor and normal genotypes and ensure concordance between samples. Cross-contamination between samples from other individuals sequenced in the same flow cell was monitored with the ContEst algorithm.(33) Samples with > 4% contamination were excluded (n = 4).

Alteration identification and annotation: The MuTect algorithm(34) was applied to identify somatic single-nucleotide variants in targeted exons. Indelocator was applied to identify small insertions or deletions(35). Gene level coverage was determined with the DepthOfCoverage in the Genome Analysis Tool Kit(36). Alterations were annotated using Oncotator(37). Power calculations for coverage were determined using the MuTect coverage file, which requires a minimum of 14X coverage in the tumor. Samples with median allelic fractions less than 0.05 and insufficient DNA for orthogonal validation with Fluidigm Access Array were excluded (n = 1). Alteration significance: MutSigCV(23) was applied to the aggregate cohort of 50 cases to determine statistically altered genes in the cohort. Alterations from all nominated significant genes from MutSigCV were manually reviewed in the Integrated Genomics Viewer (IGV)(38, 39). Alterations that were invalid based on IGV review (as a result of misalignment artifacts viewable in IGV) were subsequently excluded from the final results, which resulted in the exclusion of TGFBR1 and DEPDC4 from the final result (Supplementary Table 3).
Selective gene enrichment analysis: All somatic mutations and short insertion/deletions were aggregated for the cohort and split between responders and non-responders. Alterations with coverage at the tumor site of >= 30X and an allelic fraction >= 0.1 were considered for further analyses. Missense, nonsense, and splice site mutations, along with short insertion/deletions, were then assigned a damaging score (range 0-1) following previously reported methods(40): Missense mutations were scored using the Polyphen2 score(41) for the amino acid substitution. Missense mutations without available Polyphen2 scores (due to mapping errors or dinucleotide status) were listed as “Unavailable” in Supplementary Table 5 and excluded. Nonsense mutations, splice site mutations, and short insertion/deletions were assigned a damaging score of 1. Damaging scores are listed in Supplementary Table 5. Alterations with a damaging score ≥ 0.5 were then tabulated for occurrence in responders and non-responders. An altered gene would only be counted once per patient. Fisher’s exact test was performed to compare between cohorts to derive a p-value for each gene. Since a minimum of 6 alterations were required to observe a p-value of <= 0.01, only genes with >= 6 alterations in the cohort (thereby representing > 10% of the cohort and of highest potential clinical significance) were considered for multiple hypothesis testing. These results are summarized in Supplementary Table 6 and Fig. 2A. Comparison of ERCC2 mutation frequency in the responders was compared to the unselected TCGA and Guo et al cohorts with a binomial test. Results from this analysis are made available in Fig. 2C. Comparison of ERCC2 mutation distribution between responders and non-responders adjusted for elevated mutation rate in the responders was performed with the binomial test conditional on observing 9 mutations and using estimated ratio of mutation rates between responders and non-responders as the expected frequency of ERCC2 mutations in responders under the null hypothesis (e.g. for ERCC2:...
binom.test(x = 9, n = 9, p = Mutation_Rate_{Responders}/Mutation_Rate_{Non-Responders}, alternative = "greater").

All statistical calculations were performed using the R statistical package.

Mutation Validation

Orthogonal validation of selected mutations and short insertion/deletions (those presented in this manuscript, including ERCC2) was performed using the Fluidigm Access Array. Of 50 cases, 35 had sufficient DNA to generate sufficient read depth for analysis. A total of 85 candidate targets were submitted to Fluidigm for single-plex PCR primer assay design. This resulted in the design of 65 assays covering all 85 targets. Assay amplicons ranged from 163bp to 199bp in size, with an average of 183bp. All available samples were run on the Access Array system (Fluidigm) using three 48.48 Access Array IFC chips following manufacturer’s recommendations using the ‘4-Primer Amplicon Tagging protocol’ for Access Array (Fluidigm, P/N 100-3770, Rev. C1) with the exception that Access Array IFC chips were loaded and harvested using a Bravo Automated Liquid Handling Platform (Agilent Technologies), using manufacturer’s recommendations. Resulting amplicons containing sample specific barcodes and Illumina adapter sequencers were pooled and sequenced on a MiSeq sequencer (Illumina) with 2 runs of 150 base paired-end reads (V2 sequencing chemistry), using custom Fluidigm sequencing primers following manufacturer’s recommendations (Fluidigm). All sites were manually reviewed in IGV to determine presence or absence of non-reference reads. Details about validation results for ERCC2 and additional variants are in Supplementary Table 2. Variants where there was inadequate sample for validation or insufficient sequencing reads in the validation data to interpret manually in IGV were listed as "Unavailable".
Cloning and Cell Lines

A site-directed PCR mutagenesis/BP recombination method\(^\text{42}\) was used to generate WT and mutant ERCC2 open reading frames (ORFs). For each mutant, PCR products were generated such that fragments overlap at the region of the desired mutation. The fragments were then introduced into the pDONR vector through the BP reaction. The BP reaction mixture was transformed into \(E. \ coli\) and recombined to generate a pENTR vector. The pENTR vector was then used to perform the LR reaction to create an expression plasmid.

The expression plasmids harboring WT ERCC2, GFP (negative control), or mutant ERCC2s were expanded in \(E. \ coli\) TOP10 cells (Invitrogen) and purified using an anion exchange kit (Qiagen). Lentiviruses were propagated in 293T cells by cotransfection of the expression plasmid with plasmids encoding viral packaging and envelope proteins. Unless otherwise noted, all human cell lines were cultured in Dulbecco’s Modified Eagle’s Medium (Invitrogen) supplemented with 10% fetal bovine serum (Sigma) and 1% L-glutamine and grown at 37°C and 5% \(CO_2\). The 293T cell supernatants containing virus were collected after 48 hours, filtered twice (0.45 \(\mu\)m syringe filter, Millipore), then added directly to growing cultures of an SV40-transformed pseudodiploid ERCC2-deficient fibroblast cell line derived from an XP patient of genetic complementation group D (GM08207; Coriell Institute) The cell line is a compound heterozygote harboring ERCC2-R683W and ERCC2-DEL 36_61 mutations. Polybrene\(^\text{\textregistered}\) (Sigma) was added to a final concentration of 8 \(\mu\)g/ml to increase the efficiency of infection. Stable integrates were selected by incubation for 5 days in media containing 10 \(\mu\)g/ml blastocidin. Physical and biologic containment procedures for recombinant DNA followed institutional protocols in accordance with the National Institutes of Health Guidelines for Research.
involving Recombinant DNA Molecules. The cell line was obtained in March 2013 from Corriell Cell Repositories (Camden, NJ) and authenticated by microsatellite genotyping.

Cisplatin Sensitivity Assays

Cells were transferred to 96-well plates at a density of 500 cells per well. Six hours later, cisplatin (Sigma) was serially diluted in media and added to the wells. After 96 hours, CellTiter-Glo reagent (Promega) was added to the wells and the plates were scanned using a luminescence microplate reader (BioTek). Survival at each cisplatin concentration was plotted as a percentage of the survival in cisplatin-free media. Each data point on the graph represents the average of three independent measurements, and the error bars represent the standard deviation. IC\textsubscript{50} concentrations were calculated using a four parameter sigmoidal model and plots were generated using Prism (GraphPad). A one-way ANOVA test was used to compare the IC\textsubscript{50} of the negative control cell line to the IC\textsubscript{50} of the WT and mutant ERCC2 cell lines, and to compare the IC\textsubscript{90} of the WT line to the IC\textsubscript{50} of the negative control and mutant cell lines.

UV Clonogenic Survival Assays

Cells were seeded in 6-well plates (Nunc) at a density of 1500 cells per well. The following day, the cells were washed once with PBS and then exposed to increasing UV doses using a UV-B irradiator (Stratagene). Media was replaced and the cells were allowed to grow for nine days. On day 10, cells were fixed using a 1:5 acetic acid:methanol solution for 20 minutes at room temperature. Cells were then stained for 45 minutes using 1% crystal violet in methanol solution. Plates were rinsed vigorously with water, allowed to dry, and colonies were then manually counted. The number of colonies present at each UV dose was plotted as a ratio of the number of colonies
present in mock-irradiated wells. Each data point represents the average of three independent measurements, and the error bars represent the standard deviation.

**Chromosomal Breakage Analysis**

Approximately 1 x 10^6 cells were seeded per 10 cm dish. After 24 hours, 400 nM cisplatin was added and cells were allowed to grow for an additional 48 hours. Cells were exposed to colcemid for 2 hours, harvested using 0.075 M KCl, and fixed in 3:1 methanol:acetic acid. Slides were stained with Wright’s stain and 25-50 metaphases were analyzed for chromosomal aberrations (chromosome or chromatid breaks, rings, translocations, deletions, fragments/double minute chromosomes, tri- or quadraradials, di- or tricentrics, and premature chromatid separation).

**Immunoblots**

Frozen cell pellets were thawed and resuspended in RIPA buffer (50 mM TRIS [pH 7.3], 150 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.5% Na-deoxycholate, 0.1% SDS) supplemented with complete protease inhibitor (Roche), NaVO₄, and NaF. The cell suspensions were centrifuged and total protein concentration of the supernatant was determined by colorimetry (Bio-Rad). Samples were boiled with loading buffer (Bio-Rad) and electrophoresed in a 3-8% gradient TRIS-acetate gel (Life Technologies). Resolved proteins were transferred to a PVDF membrane (Millipore) at 90V for 2 hours at 4°C. The membrane was blocked for one hour in blocking solution (5% milk in TRIS-buffered saline-T) and incubated with primary antibody in blocking solution at 4°C overnight (mouse ERCC2, AbCam; rabbit β-actin, Cell Signaling). The following day, the membrane was rinsed and incubated for one hour with peroxidase-conjugated secondary antibody in blocking solution (anti-mouse and anti-rabbit, Cell Signaling) and
rinsed. Enhanced chemiluminescent substrate solution (PerkinElmer) was added and signal was detected by film exposure (GE Healthcare).

ACKNOWLEDGEMENTS

The authors wish to acknowledge the patients for contributing tissue for research. In addition, we thank Dr. Kenna Shaw, Dr. Jean C. Zenklusen, and The Cancer Genome Atlas for allowing pertinent ERCC2 results to be incorporated into our analyses. We also thank the Broad Genomics Platform for sequencing and validation activities, and Lisa Moreau for help with chromosomal studies.
REFERENCES


26. Cleaver JE, Thompson LH, Richardson AS, States JC. A summary of mutations in the UV-sensitive disorders: xeroderma pigmentosum, Cockayne syndrome, and...


### TABLES

**Table 1. Patient characteristics.** Clinical characteristics of the total patient cohort, as well as data stratified by responder or non-responder status. *For patients alive at the time of this study only. P < 0.05 is considered significant (Mann-Whitney Test). (TUR:}

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Total (N = 50)</th>
<th>Responder (N = 25)</th>
<th>Non-Responders (N = 25)</th>
<th>p &gt; 0.05</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Age at TUR - yr</strong></td>
<td>62.5 ± 8.9</td>
<td>61 ± 10.1</td>
<td>64 ± 7.3</td>
<td>p &gt; 0.05</td>
</tr>
<tr>
<td><strong>Female sex - no. (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hispanic/Latino</td>
<td>1 (2)</td>
<td>0 (0)</td>
<td>1 (4)</td>
<td></td>
</tr>
<tr>
<td>Non-Hispanic/Non-Latino</td>
<td>48 (96)</td>
<td>24 (96)</td>
<td>24 (96)</td>
<td></td>
</tr>
<tr>
<td>Unknown</td>
<td>1 (2)</td>
<td>1 (4)</td>
<td>0 (0)</td>
<td></td>
</tr>
<tr>
<td><strong>Race - no. (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Caucasian</td>
<td>49 (98)</td>
<td>25 (100)</td>
<td>24 (96)</td>
<td></td>
</tr>
<tr>
<td>African-American</td>
<td>1 (2)</td>
<td>0 (0)</td>
<td>1 (4)</td>
<td></td>
</tr>
<tr>
<td><strong>Smoking Status - no. (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Never</td>
<td>13 (26)</td>
<td>8 (32)</td>
<td>5 (20)</td>
<td></td>
</tr>
<tr>
<td>Former</td>
<td>26 (52)</td>
<td>13 (52)</td>
<td>13 (52)</td>
<td></td>
</tr>
<tr>
<td>Current</td>
<td>11 (22)</td>
<td>4 (16)</td>
<td>7 (28)</td>
<td></td>
</tr>
<tr>
<td><strong>Clinical Staging - no. (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T2</td>
<td>37 (74)</td>
<td>19 (76)</td>
<td>18 (72)</td>
<td></td>
</tr>
<tr>
<td>T3</td>
<td>10 (20)</td>
<td>5 (20)</td>
<td>5 (20)</td>
<td></td>
</tr>
<tr>
<td>T4</td>
<td>3 (6)</td>
<td>1 (4)</td>
<td>2 (8)</td>
<td></td>
</tr>
<tr>
<td>N0</td>
<td>40 (80)</td>
<td>18 (72)</td>
<td>22 (88)</td>
<td></td>
</tr>
<tr>
<td>Node Positive</td>
<td>10 (20)</td>
<td>7 (28)</td>
<td>3 (12)</td>
<td></td>
</tr>
<tr>
<td><strong>TUR Histology - no. (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TCC</td>
<td>32 (64)</td>
<td>19 (76)</td>
<td>13 (52)</td>
<td></td>
</tr>
<tr>
<td>Mixed TCC</td>
<td>18 (36)</td>
<td>6 (24)</td>
<td>12 (48)</td>
<td></td>
</tr>
<tr>
<td><strong>Neoadjuvant Chemotherapy Regimen - no. (%)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GC</td>
<td>31 (62)</td>
<td>14 (56)</td>
<td>17 (68)</td>
<td></td>
</tr>
<tr>
<td>ddMVAC</td>
<td>16 (32)</td>
<td>9 (36)</td>
<td>7 (28)</td>
<td></td>
</tr>
<tr>
<td>GC-Sunitinib</td>
<td>2 (4)</td>
<td>2 (8)</td>
<td>0 (0)</td>
<td></td>
</tr>
<tr>
<td>ddGC</td>
<td>1 (2)</td>
<td>0 (0)</td>
<td>1 (4)</td>
<td></td>
</tr>
<tr>
<td><strong>Median interval from chemotherapy to cystectomy (Days ±SD)</strong></td>
<td>47 ± 29.9</td>
<td>46 ± 26.0</td>
<td>47 ± 33.8</td>
<td>p &gt; 0.05</td>
</tr>
<tr>
<td><strong>Median length of follow-up (Days ±SD)</strong></td>
<td>351 ± 372.5 ± 329.5 ± 287.1</td>
<td>p &gt; 0.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Patients with Recurrence - no. (%)</strong></td>
<td>13 (26)</td>
<td>2 (8)</td>
<td>11 (44)</td>
<td></td>
</tr>
</tbody>
</table>
transurethral resection; GC: gemcitabine and cisplatin; ddMVAC: dose dense methotrexate, vinblastine, doxorubicin, and cisplatin; ddGC: dose dense gemcitabine and cisplatin; SD: standard deviation). Plus-minus values are medians ± standard deviation.
FIGURE LEGENDS

Figure 1. Study design, mutation rates, and aggregate significant somatic mutations. Panel A shows patients with muscle-invasive urothelial carcinoma cancer split into cases and controls based on their pathologic response to cisplatin-based neoadjuvant chemotherapy (TURBT: transurethral resection of bladder tumor). Nine cases could not complete sequencing due to technical reasons (failed sequencing or elevated contamination). Data in Panel B are arranged so that each column represents a tumor and each row represents a gene. The center panel is divided into responders (left and black) and non-responders (right and yellow). The mutation rates of responders are elevated compared to non-responders (top of Panel B). The alteration landscape (center of Panel B) of the aggregate cohort (n = 50 patients) demonstrates a set of statistically significant genes that are altered in urothelial carcinoma (TP53, RB1, KDM6A, ARID1A). The negative log of the q values for the significance level of mutated genes is shown (for all genes with q < 0.1) on the right side of Panel B. ERCC2 mutation status is also shown below the other genes, although ERCC2 was not significantly mutated across the combined cohort. Additional data regarding allelic fraction ranges for each case (bottom of Panel B), mutation rates (top of Panel B), and mutational frequency (left of Panel B) are also summarized in this figure.

Figure 2. Three tests examining selective enrichment of ERCC2 mutations in cisplatin-responder tumors. Panel A shows a plot of MutSigCV gene-level significance (-log10(MutSigCV p-value) and responder enrichment significance (-log10(Fisher’s exact test p-value)). The size of the point is proportional to the number of responder patients who harbor alterations in the gene. Genes with a responder enrichment p-value of < 0.01 are colored red; others are colored gray, and the dashed line denotes a p value of 0.01. Only ERCC2 reaches statistical significance in the responder cohort (P < 0.001;
Fisher’s exact test). In Panel B, among genes with sufficient number of alterations for cohort comparisons \((n = 9)\), only ERCC2 somatic mutations occur exclusively in the cisplatin responders, which is significant when accounting for the elevated mutation rate in responders compared to non-responders \((P < 0.05, \text{denoted by asterisk})\). Compared to unselected TCGA and Guo et al urothelial carcinoma cohorts, Panel C shows that ERCC2 somatic mutations are significantly enriched in the responder cohort \((P < 0.01, \text{denoted by asterisk})\).

**Figure 3.** ERCC2 mutation mapping and distribution across tumor types. Panel A depicts a stick plot of ERCC2 showing the locations of somatic mutations in the responders compared to ERCC2 mutations observed in two separate unselected bladder cancer exome cohorts. The ERCC2 mutations cluster within or near conserved helicase motifs. Panel B illustrates the somatic ERCC2 mutation frequency in multiple tumor types from the Cancer Genome Atlas (TCGA). In Panel C, the structure of an archaebacterial ERCC2 (PDB code: 3CRV) with mutations identified in the responder cohort mapped to their equivalent position is illustrated. These locations are shown in the context of canonical germline ERCC2 mutations responsible for xeroderma pigmentosum (XP), xeroderma pigmentosum/Cockayne Syndrome (XP/CS), and trichothiodystrophy (TTD).

**Figure 4.** ERCC2 mutants fail to rescue cisplatin sensitivity of ERCC2-deficient cells. Panel A shows an immunoblot of ERCC2 expression in cell lines created by transfection of the ERCC2-deficient parent cell line (GM08207; Coriell Institute) with pLX304 (Addgene) encoding GFP (negative control), WT ERCC2, or a mutant ERCC2. The negative control ERCC2-deficient cell line (lane 1) expresses endogenous levels of inactive ERCC2 from the parent cell genome, whereas WT (lane 2) and mutant (lanes 3-7) ERCC2 cell lines show increased levels of ERCC2 expressed from the transfected
gene. β-actin is shown as a loading control. Panel B shows the cisplatin sensitivity profiles of cell lines expressing WT or mutant ERCC2. Expression of WT ERCC2 in an ERCC2-deficient background rescues cisplatin sensitivity, whereas expression of the ERCC2 mutants fails to rescue cisplatin sensitivity. An IC$_{50}$ was calculated from the survival data for each cell line and these values are shown in Panel C. The difference in IC$_{50}$ between the parent (ERCC2-deficient) cell line and the cell line expressing WT ERCC2 was statistically significant, as was the difference between the WT ERCC2 cell line and each of the mutant ERCC2 cell lines ($P < 0.0001$; ANOVA). The difference between the ERCC2-deficient cell line and each of the mutant cell lines was not statistically significant.

**Figure 5. ERCC2 mutants fail to rescue UV sensitivity of ERCC2-deficient cells.** Panel A shows a representative colony formation assay for the ERCC2-deficient cell line (top) as well as the ERCC2-deficient line transfected with WT ERCC2 (middle) or one of the ERCC2 mutants (D609G, bottom) following increases doses of UV irradiation. Panel B shows clonogenic survival data for negative control, WT, and mutant ERCC2 cell lines. WT ERCC2 rescues UV sensitivity of the ERCC2-deficient cell line whereas the mutant ERCC2s fail to rescue UV sensitivity. In Panel C, UV IC$_{50}$ values for cell lines are shown. The difference between the ERCC2-deficient cell line and the WT ERCC2 cell line was significant ($P < 0.0001$; ANOVA), whereas the difference between the ERCC2-deficient cell line and each of the ERCC2 mutant cell lines was not statistically significant (NS).

**Figure 6. ERCC2 mutants fail to rescue genomic instability following cisplatin exposure.** Representative mitotic spreads from an ERCC2-deficient cell line (Panel A), and the same ERC2-deficient cell line transfected with WT ERCC2 (Panel B) or one of the ERCC2 mutants (V242F, Panel C) following cisplatin exposure. Panel D shows chromosomal aberration data from ERCC2-deficient, WT ERCC2, and mutant ERCC2
cell lines. Rates of chromosomal aberrations following cisplatin exposure were significantly lower in the WT ERCC2 cell line than in the ERCC2-deficient line or the cell lines expressing mutant ERCC2 ($P = 0.03$; ANOVA)
Figure 2

A

Responder Significance

MutSigCV Significance

B

Responders (pT0/Tis)  Non-responders (pT2+)

* ERCC2

NF1

PIK3CA

ERBB4

RB1

FAT4

ARID1A

KDM6A

TP53

C

Patients with somatic ERCC2 mutations (%)

- Responders
- Unselected

0 10 20 30 40

pT0/Tis TCGA Guo et al (n = 99)

Number of patients

15
Figure 4

A)

1. ERCC2\(^{-/-}\) mock
2. ERCC2\(^{-/-}\) WT
3. ERCC2\(^{-/-}\) V242F
4. ERCC2\(^{-/-}\) P463L
5. ERCC2\(^{-/-}\) E606G
6. ERCC2\(^{-/-}\) D609G
7. ERCC2\(^{-/-}\) G665A

B)

Relative Cell Viability

C)

\( p = \text{NS} \)

\( p < 0.0001 \)
Figure 5

A

UV Dose (J/m²)

ERCC2⁺⁻ mock

ERCC2⁺⁻ WT

ERCC2⁺⁻ D609G

B

Survival Ratio

UV Dose (J/m²)

ERCC2⁺⁻ mock

ERCC2⁺⁻ WT

ERCC2⁺⁻ V242F

ERCC2⁺⁻ P463L

ERCC2⁺⁻ E606G

ERCC2⁺⁻ D609G

ERCC2⁺⁻ G665A

C

p = NS

p < 0.0001

IC₅₀ (J/m²)

UV Dose (J/m²)
Somatic ERCC2 mutations correlate with cisplatin sensitivity in muscle-invasive urothelial carcinoma

Eliezer M Van Allen, Kent W. Mouw, Philip Kim, et al.

Cancer Discovery Published OnlineFirst August 5, 2014.