

SCIENTIFIC REPORTS



OPEN

Genomic features of renal cell carcinoma with venous tumor thrombus

Gregor Warsaw¹, Daniel Hübschmann^{1,2,3}, Kortine Kleinheinz^{1,2}, Cathleen Nientiedt^{4,5,6}, Martina Heller⁴, Laura Van Coile⁴, Yanis Tolstov⁴, Lukas Trennheuser⁷, Kathrin Wiczorek⁸, Carine Pecqueux^{6,7}, Claudia Gasch^{6,7}, Timur Kuru^{7,12}, Joanne Nyarangi-Dix^{6,7}, Gencay Hatiboglu^{6,7}, Dogu Teber^{6,7}, Sven Perner⁹, Albrecht Stenzinger⁸, Wilfried Roth^{8,13}, Boris Hadaschik^{6,7,14}, Sascha Pahernik^{6,7,15}, Dirk Jäger^{5,6}, Carsten Grüllich^{5,6}, Anette Duensing¹⁰, Roland Eils^{1,2}, Matthias Schlesner^{1,16}, Holger Sültmann¹¹, Markus Hohenfellner^{6,7} & Stefan Duensing^{4,6,7}

A venous tumor thrombus (VTT) is a potentially lethal complication of renal cell carcinoma (RCC) but virtually nothing is known about the underlying natural history. Based on our observation that venous thrombi contain significant numbers of viable tumor cells, we applied multiregion whole exome sequencing to a total of 37 primary tumor and VTT samples including normal tissue specimens from five consecutive patients. Our findings demonstrate mutational heterogeneity between primary tumor and VTT with 106 of 483 genes (22%) harboring functional SNVs and/or indels altered in either primary tumor or thrombus. Reconstruction of the clonal phylogeny showed clustering of tumor samples and VTT samples, respectively, in the majority of tumors. However, no new subclones were detected suggesting that pre-existing subclones of the primary tumor drive VTT formation. Importantly, we found several lines of evidence for “BRCAness” in a subset of tumors. These included mutations in genes that confer “BRCAness”, a mutational signature and an increase of small indels. Re-analysis of SNV calls from the TCGA KIRC-US cohort confirmed a high frequency of the “BRCAness” mutational signature AC3 in clear cell RCC. Our findings warrant further pre-clinical experiments and may lead to novel personalized therapies for RCC patients.

¹Division of Theoretical Bioinformatics (B080), German Cancer Research Center (DKFZ), Im Neuenheimer Feld 280, D-69120, Heidelberg, Germany. ²Department for Bioinformatics and Functional Genomics, Institute for Pharmacy and Molecular Biotechnology (IPMB) and BioQuant, Heidelberg University, Im Neuenheimer Feld 267, D-69120, Heidelberg, Germany. ³Department of Pediatric Immunology, Hematology and Oncology, University Hospital Heidelberg, Im Neuenheimer Feld 430, D-69120, Heidelberg, Germany. ⁴Section of Molecular Urooncology, Department of Urology, University of Heidelberg School of Medicine, Im Neuenheimer Feld 517, D-69120, Heidelberg, Germany. ⁵Department of Medical Oncology, University of Heidelberg School of Medicine, National Center for Tumor Diseases (NCT), Im Neuenheimer Feld 460, D-69120, Heidelberg, Germany. ⁶Center for Kidney Tumors, National Center for Tumor Diseases and University of Heidelberg School of Medicine, Im Neuenheimer Feld 460, D-69120, Heidelberg, Germany. ⁷Department of Urology, University of Heidelberg School of Medicine, Im Neuenheimer Feld 110, D-69120, Heidelberg, Germany. ⁸Institute of Pathology, University of Heidelberg School of Medicine, Im Neuenheimer Feld 224, D-69120, Heidelberg, Germany. ⁹Pathology of the University Hospital Schleswig-Holstein, Campus Lübeck and the Research Center Borstel, Leibniz Lung Center, Ratzeburger Allee 160, D-23538 Lübeck and Parkallee 1-40, D-23845, Borstel, Germany. ¹⁰University of Pittsburgh Cancer Institute, Cancer Therapeutics Program, 5117 Centre Avenue, Pittsburgh, PA 15213, USA. ¹¹National Center for Tumor Diseases, German Cancer Research Center, Cancer Genome Research, Im Neuenheimer Feld 460 and German Cancer Consortium (DKTK), D-69120, Heidelberg, Germany. ¹²Present address: Department of Urology, University of Cologne, Kerpener Str. 62, D-50937, Cologne, Germany. ¹³Present address: Institute of Pathology, University of Mainz Medical School, Langenbeckstr. 1, D-55131, Mainz, Germany. ¹⁴Present address: Department of Urology, University Hospital Essen, University of Duisburg-Essen, Hufelandstr. 55, D-45122, Essen, Germany. ¹⁵Present address: Department of Urology, Nuremberg Hospital, Paracelsus Medical University, Prof.-Ernst-Nathan-Strasse 1, D-90419, Nuremberg, Germany. ¹⁶Present address: Bioinformatics and Omics Data Analytics (B240), German Cancer Research Center (DKFZ), Im Neuenheimer Feld 280, D-69120, Heidelberg, Germany. Gregor Warsaw and Daniel Hübschmann contributed equally to this work. Matthias Schlesner, Holger Sültmann, Markus Hohenfellner and Stefan Duensing jointly supervised this work. Correspondence and requests for materials should be addressed to S.D. (email: stefan.duensing@med.uni-heidelberg.de)

Metastatic renal cell carcinoma (RCC) is the most lethal urological cancer and characterized by extensive genomic intratumoral heterogeneity^{1,2}. In approximately 4–25% of RCC patients, the primary tumor extends into the renal vein, the inferior vena cava or as far as the right atrium of the heart, usually under formation of a large blood clot^{3–5}. RCC tumor thrombi are subdivided macroscopically according to their consistency into solid or friable. The latter has been suggested to be associated with a worse prognosis although this notion has recently been challenged^{6–8}. Despite the poor prognosis associated with venous tumor thrombus (VTT) formation if left untreated^{9–11}, virtually nothing is known about its natural history.

Clear cell RCC has previously been shown to exhibit a high degree of mutational heterogeneity and patterns of branched and convergent clonal evolution^{1,2}. There is evidence that functional and genomic intratumoral heterogeneity in RCC is shaped by spatial niches^{12,13}. To what extent these findings apply to RCC with VTT is currently not known in detail¹⁴.

Herein, we used multiregion whole exome sequencing of RCCs with VTT to characterize the mutational landscape of these tumors. We found heterogeneous mutations between primary tumor and VTT, however, our results suggest that pre-existing subclones of the primary tumor drive VTT formation. Importantly, we provide evidence for homologous recombination defects (“BRCAness”) in a subset of RCCs with VTT, which holds the promise for novel personalized therapeutic interventions in patients with this condition including PARP inhibitors, platinum salts or others¹⁵.

Methods

Patients. A total of five consecutive patients with RCC and venous tumor thrombus were prospectively included after written informed consent and under Ethics committee vote 207/2005 of the University of Heidelberg School of Medicine. All patients showed a clear cell histomorphology except patient RCC-VTT-04 with a papillary type II RCC.

Whole exome sequencing. DNA extraction and whole exome sequencing were performed by GATC Biotech (Konstanz, Germany) on an Illumina platform (InView™ Human Exome). Fastq files were aligned against the human reference genome (build 37, version hs37d5), using bwa mem (version 0.7.8 with minimum base quality threshold set to zero [–T 0] and remaining settings left at default values), followed by coordinate-sorting with bamsort (with compression option set to fast (1)) and marking duplicate read pairs with bammarkduplicates (with compression option set to best (9)) (both part of the biobambam package version 0.0.148).

The generated bam files were used to identify high-confidence somatic SNVs by applying an in-house analysis pipeline based on samtools¹⁶ and bcftools (both version 0.1.19) as described earlier^{17,18}. Briefly, samtools version 0.1.19¹⁶ was used for calling all candidate variants (positions where the tumor sample differs from the reference sequence). The filter, which usually makes samtools a germline caller, was switched off. Candidate variants were postfiltered by custom filters optimized for somatic variant calling. The discrimination between somatic and germline variants was performed by looking up the variant position in the matched normal sample.

Results were amended by additional classification of the SNVs into 96 categories according to their triplet of the base preceding the SNV, the actual transition/transversion and the following base. For each class it was determined whether its SNVs exhibit a sequencing strand bias. For classes with bias, SNVs with less than two supporting reads on the strand opposite to the strand exhibiting the biased variants were filtered out. Somatic SNVs were called functional, when they were nonsynonymous, occurred within splice sites, or lead to gain or loss of stop codons. All SNVs were annotated using Annovar¹⁹ (version as of February 2016), dbSNP (build 147)²⁰, 1000 Genomes²¹ and EVS (Exome Variant Server, NHLBI GO Exome Sequencing Project (ESP), Seattle, WA [version ESP6500SI-V2]) data. Indels were identified with Platypus callVariants (version 0.8.1, parameter settings genIndels = 1, genSNPs = 1, bufferSize = 100000 and maxReads = 1000000)²². All indels were annotated using Annovar¹⁹ (version as of November 2014), dbSNP (build 141)²⁰, 1000 Genomes²¹ data.

Sequence data for this study is available in the European Genome-phenome Archive (EGA), which is hosted by the EBI, under accession number EGAS#00001001950.

Supervised Analysis of Mutational Signatures. A supervised analysis of mutational signatures determines contributions of known mutational signatures to the mutational catalogue, i.e. counts of nucleotide exchanges in their motif context, of a given dataset. As opposed to an unsupervised analysis of mutational signatures, i.e. a discovery method like non-negative matrix factorization (NMF) as used in²³, a supervised analysis requires much less statistical power, but will not discover new signatures. A software to carry out such an analysis was written as an R package (YAPSA – *Yet Another Package for Signature Analysis*²⁴, manuscript in preparation). Due to the low requirements on statistical power, i.e. the number of samples in a cohort and the number of point mutations per sample, the supervised analysis of mutational signatures as carried out by YAPSA may be run on very small cohorts.

Using YAPSA, a linear combination decomposition of the mutational catalogue with known and predefined signatures was computed by non-negative least squares (NNLS), functions for which are implemented in the R package *lsei*²⁵. In order to increase specificity, the NNLS algorithm was applied twice: after the first execution, only those signatures whose exposures, i.e., contributions in the linear combination, were higher than a certain cutoff, were kept and the NNLS was run again with the reduced set of signatures. As detectability of the different signatures may vary, the cutoffs were chosen to be signature-specific.

In the analysis presented here, we used a set of 30 publicly available mutational signatures (<http://cancer.sanger.ac.uk/cosmic/signatures>), which were identified in an unsupervised analysis of 10,952 whole exome sequenced and 1,048 whole genome sequenced tumor normal control pairs (12,000 pairs in total). We call these signatures, which are an extension of the signatures published initially in²³, AC1 – AC30 (AC standing for *Alexandrov COSMIC*). The signature-specific cutoffs were determined in a receiver operator characteristic (ROC)

analysis using publicly available data on mutational catalogues of the samples included in the initial Alexandrov analysis²³, consisting of 7,042 tumor normal control pairs (507 from whole genome sequencing and 6,535 from whole exome sequencing).

Before running an analysis, YAPSA builds up the mutational catalogues of the samples or cohorts to be analyzed by calling the functions *mutationContext* and *mutationContextMatrix* from the R package *SomaticSignatures*²⁶. The mutational catalogue was then corrected for different occurrences of the motifs between the whole genome and the target capture used for whole exome sequencing of the tumor of our patients.

Copy number alterations (CNAs), tumor cell content (TCC) and ploidy estimation. CNAs were inferred from whole exome sequencing data with *cnvKit* version 0.8.6 (git repository hash: 85774ac)²⁷ with default parameter settings.

Heterozygous SNPs were determined as those positions with alternative allele fraction between 0.3 and 0.7 in the respective normal sample. Segments which contained at least 20 heterozygous SNPs were further processed to infer sample ploidy and tumor cell content along with allele-specific copy number estimates. The segments were classified as balanced or imbalanced according to the distribution of the frequencies of the alternative allele of the SNPs in the respective segments. If this distribution had a global maximum between 0.45–0.55 a segment was called balanced, remaining segments were further separated into two groups – ambiguous segments with one density peak outside of the above mentioned interval, and imbalanced segments with two peaks. Ambiguous segments were neglected in subsequent steps. For imbalanced segments, the mean B-allele frequency (BAF) of all SNPs in the segment that were heterozygous in the germline was estimated using the allele with the higher read count as B-allele. Then the mean read count of the B-allele was calculated as product of total coverage and the BAF of the respective segment.

Tumor cell content (TCC) and ploidy of a sample were estimated using a method adapted from ACEseq (manuscript in preparation). For TCC estimation values in the range of 0.15–1.0 were tested whereas a ploidy range between 1 and 6.5 was allowed. For each possible combination of TCC and ploidy, absolute copy numbers as well as allele-specific copy numbers and the decrease in heterozygosity (DH)²⁸ were estimated segment-wise. Allele-specific copy numbers were calculated as total copy number divided by two for balanced segments and as a function of coverage and B-allele read counts in case of imbalanced segments. The weighted mean distance of all segments to the next allowed integer copy number state was calculated for total and allele-specific copy numbers. Here, allowed means even total copy number states for balanced segments and any integer copy number state for imbalanced segments and allele-specific copy numbers. TCC/ploidy combinations requiring negative copy number states or a DH larger than 1 for any segment were excluded. Local minima in the weighted mean distance were considered as possible TCC/ploidy solution for the sample and were visually evaluated. Additionally, TCC was estimated from the mutant allele fraction distribution of somatic SNVs and CNV-based and SNV-based estimates were compared.

Due to low TCC, resulting in a weak signal for imbalance detection, some samples required manual adjustments. For the sample RCC-VTT-01 thrombus III (7), regions that were erroneously annotated as balanced by the automated detection were reclassified as imbalanced based on the coverage profile and comparison to the remaining samples from the same patient. B-allele read counts were recalculated for these segments and a rerun of ploidy and TCC estimation was performed. All samples from patients RCC-VTT-02 and RCC-VTT-03 showed a low TCC. Only one sample from RCC-VTT-02 (tumor (1)) had sufficient TCC for proper calling of allelic states. For all other RCC-VTT-02 samples the restrictions for balanced segments were loosened so they could be fitted to uneven copy numbers too to account for undetected imbalances. This led to a better fit of the copy numbers though TCC was still estimated too high based on visual inspection of the BAF profiles and compared to the MAF estimates, most likely due to high subclonality or low signal due to the low TCC. RCC-VTT-02 tumor (1) was used as reference for the newly obtained calls to evaluate the allele-specific estimates as some of its CNAs were also expected to appear in the other samples with lower purity. The new ploidy fits were combined with TCCs that were based on the MAF estimates. Good results could be obtained with this approach with respect to small deviations from integer copy number states and overall comparison against RCC-VTT-02 tumor (1). For patient RCC-VTT-03, all samples showed low TCC (<30%, based on MAF estimates and visual inspection of BAF profiles), so no reliable estimates for balanced and imbalanced states could be achieved. As we were missing a reference for this case, no allele specific copy numbers could be obtained for this patient. Similar to RCC-VTT-02, TCC was based on the MAF distribution whereas the ploidies were estimated with the constraint that the lowest observed coverage had to correspond to one copy and the majority of the genome fitted to a diploid state.

HRD (Homologous Recombination Deficiency) and LST (Large Scale Transition) scores estimation.

For stable HRD and LST scores estimation, copy number profiles were smoothed to reduce oversegmentation caused by technical noise. All neighboring segments that rounded to the same total and allele specific copy number and did not deviate by more than 0.3 from another were merged. In addition to smoothing of similar segments any segments smaller than 3 Mb were merged with their more similar neighbor as suggested by Popova *et al.*²⁹.

Any switch between copy number states of segments that were larger than 10 Mb and did not correspond to full chromosome arms was counted as LST according to Popova *et al.*²⁹. Additionally, segments that were larger than 15 Mb but less than a whole chromosome in length and corresponded to a loss of heterozygosity (LOH) were counted for the HRD estimation³⁰.

In case of RCC-VTT-03, only segments with a single copy were taken into account for HRD estimation due to the lack of allele-specific copy numbers.

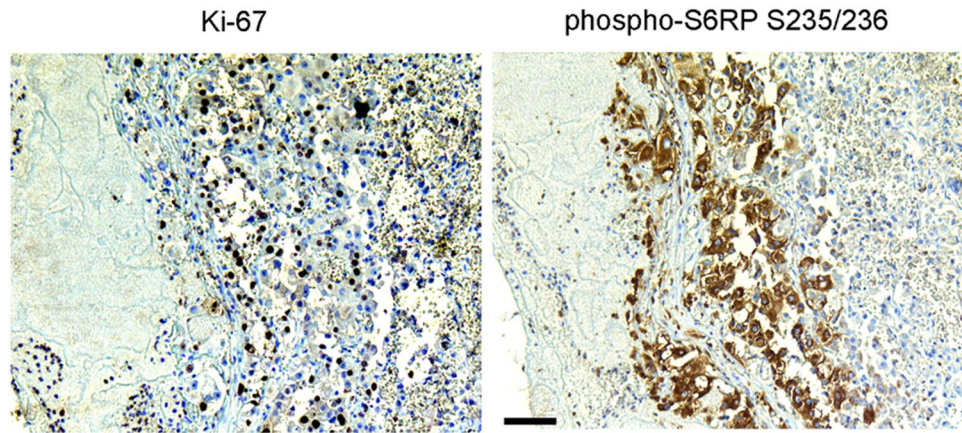


Figure 1. Tumor thrombi contain vital tumor cells. Representative immunohistochemical staining of a venous tumor thrombus for Ki-67 and phospho-S6RP S235/236. The tumor was unrelated to the five RCC-VTT cases analyzed. Scale bar = 100 μ m.

Heterogeneity analysis. In order to reconstruct the clonal evolution of the tumor samples, the R package Canopy version 1.1.1 was used³¹. Canopy takes two sources of information as input: MAF values of SNVs and CNAs. In this project, functional somatic SNVs and smoothed copy number segments (as described above) were used. Segments, which were shared by all samples were removed, as they do not contribute to discrimination between samples. Additionally, according to the advice of the author of the software, small segments (<1 kbp) were neglected, as larger events tend to be more reliable than small ones, especially on exome sequencing data. For each patient, the optimal number of subclones K was identified by Canopy based on the Bayes Information Criterion (BIC), followed by a Markov chain Monte Carlo (MCMC) based approach to determine a tree topology with K leaves and a mutation setting in this tree, which explains the observed mutation pattern with highest likelihood. MCMC sampling steps were performed over 50,000 iterations in 30 individual chains (to avoid local optima) in a parallelized manner. Mutations defining individual subclones are available upon request.

Immunohistochemistry. Formalin-fixed, paraffin-embedded tissue specimens were provided by the tissue bank of the National Center for Tumor Diseases (NCT, Heidelberg, Germany) in accordance with the regulations of the tissue bank and the approval of the ethics committee of the University of Heidelberg School of Medicine. Paraffin sections were deparaffinized in xylene and rehydrated in a graded ethanol series. Antigen retrieval was performed with a steam cooker using retrieval buffer (Target Retrieval Solution, Dako). Primary antibodies used were directed against Ki-67 (clone MIB-1, Dako, 1:100), phospho-S6 ribosomal protein S235/236 (Cell Signaling; 1:100) or γ H2AX (clone JBW301, Millipore, 1:100) and were incubated overnight at 4 °C. Immunodetection was performed using the Histostain-Plus Detection Kit (3rd Generation, Invitrogen) according to manufacturer's recommendations. Nuclear counterstaining was provided by hematoxylin (Thermo Scientific). Tissue specimens were analyzed by two independent observers (L.V.C., M.He.). For statistical analysis, the Mann Whitney U test, Wilcoxon rank-sum test or Student's t test for independent samples were used wherever applicable. Differences with a p value of ≤ 0.05 were considered significant.

All methods were performed in accordance with the relevant guidelines and regulations and experimental protocols were approved by the Ethics committee of the University of Heidelberg School of Medicine (206/2005). Datasets generated during the current study and not deposited at EGA are available from the corresponding author at reasonable request.

Somatic variant calls in the TCGA KIRC-US cohort. Somatic SNV calls from the TCGA KIRC-US cohort³² were downloaded from <https://portal.gdc.cancer.gov/legacy-archive/files/3e93f2c1-559f-4747-9c53-30589dbc03f7> as per February 5th, 2018. Samples were filtered for mutational load and only those with more than 25 somatic SNVs were kept for the analysis, leaving 400 samples for the analysis. In complete analogy to the analysis of mutational signatures of our own samples, a mutational catalog was built, correction for triplet content was performed and a supervised analysis of mutational signatures using YAPSA²⁴ was performed.

Results

Venous tumor thrombi contain viable and proliferating RCC cells. In exploratory experiments that preceded the multiregion whole exome sequencing study, we analyzed representative samples from 14 RCCs with VTT to assess if viable cells were present in the blood clot (Fig. 1). Unexpectedly, we detected a considerable number of proliferating Ki-67-positive tumor cells in the venous thrombi (mean 25.6 cells per 40 \times high-power field [HPF]; range, 2.3–88.7 cells/HPF). Moreover, we found tumor cells showing activated growth and survival signaling via the PI3K-AKT-mTOR axis as evidenced by positivity for phospho-S6RP S235/236 (mean 4.8 cells per HPF; range 0–11.9 cells/HPF).

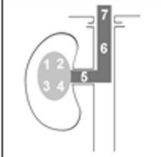

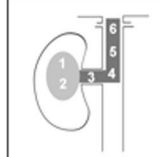
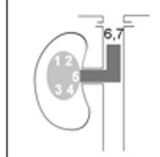
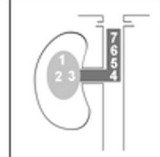
Patient	RCC-VTT-1	RCC-VTT-02	RCC-VTT-03	RCC-VTT-04	RCC-VTT-05
Age (years)	59	47	79	83	51
Sex (M/F)	F	M	M	M	F
TNM stage	pT3cpN0cM1	pT3bpNxcM0	pT3bpN0cM0	pT3bpN2cM1	pT3bpN1cM0
ISUP grade	2	2	2	3	3
Thrombus level	IV	I	III	II	II
Consistency	solid	solid	solid	friable	solid
Histology	clear cell	clear cell	clear cell	papillary type II	clear cell
Sampling sites					

Figure 2. Patient characteristics and sampling sites. VTT levels are indicated according to Novick⁴⁸.

Mutational landscape of RCC with VTT. Encouraged by the finding that venous thrombi contain a substantial amount of viable tumor cells, we performed multiregion whole exome sequencing of a total of 37 samples from five consecutive patients with RCC and VTT (Fig. 2) to directly compare the mutational landscape between the primary tumor and the corresponding VTT. Four patients were diagnosed with clear cell renal cell carcinoma whereas one patient (RCC-VTT-04) had a poorly differentiated type II papillary RCC and, in line with previous reports, mutations in *NF2* and *SETD2* (see below)^{33,34}. The latter patient had a friable tumor thrombus, whereas the other four were of solid consistency (Fig. 2).

DNA isolated from normal tissue ($n = 5$), different locations of the primary tumors ($n = 16$) and different levels of venous tumor thrombi ($n = 16$; Fig. 2) were sequenced to >35 -fold target coverage (mean target coverage: $48.7 \times$ [range: $37 \times - 99.5 \times$]; Suppl. Table 1). In all patients, the minor allele frequency (MAF) distribution was comparable between samples from the primary tumor and VTT indicating that results were not due to an uneven admixture of non-tumorous cells (Suppl. Fig. 1).

Across all samples, we identified 442 somatic functional SNVs and 64 functional indels in a total of 483 genes (Suppl. Table 2). Known RCC-associated genes and drivers were mostly found to be mutated in both the primary tumor and the VTT (*VHL*, *PBRM1*, *SETD2*, *BAP1*, *PTEN*, *MTOR* and *TP53*) with the exception of an *MTOR* mutation in RCC-VTT-02 and an *ARID1A* mutation in RCC-VTT-05 (Fig. 3)³². Among the genes mutated, a number of DNA damage checkpoint and repair genes stood out including *BAP1*, *CDK12*, *CEP164*, *ERCC1*, *LIG1*, *POLK*, *POLM* and *USP28* (Fig. 3)³⁵⁻³⁷.

Overall, 106 of 483 genes (22%) harboring functional SNVs and/or indels were heterogeneous between primary tumor and thrombus i.e., they were altered only in one but not the other of the two compartments (Suppl. Table 2 and Fig. 3). The sum of the number of genes that were mutated exclusively in the primary tumor (by somatic SNVs and/or indels) but not the VTT of the same patient was 63 (13% of all altered genes; patient-individual numbers ranged from 1 to 24) across all patients. Conversely, a total of 43 genes (8.9% of all altered genes) were mutated exclusively in venous tumor thrombi but not in the corresponding primary tumor (range from 5 to 13; Suppl. Table 2).

Copy number alterations (CNAs) inferred from exome sequencing data after correction for purity showed a loss of chromosome 3p except in patient RCC-VTT-04 with a papillary type II RCC (Suppl. Fig. 2).

Taken together, these results underscore that venous thrombi in RCC contain viable tumor cells with heterogeneous mutations, some of which affect genes involved in the cellular response to DNA damage.

Clonal phylogeny in RCC with VTT. We next sought to analyze the relationship between primary tumor and VTT in greater detail. To reconstruct the phylogeny of the tumor samples, the R package Canopy was used³¹. With the help of Canopy, we attempted to reconstruct the subclonal evolution of the tumor and subclone contributions to the individual primary tumor and VTT samples. Canopy was applied to the patients individually. As input for Canopy, we took the smoothed copy number segment data and the collection of all SNVs which were identified as somatic and functional in any of the samples of the respective patient. As shown in Fig. 4, this approach was able to model the clonal evolution and waves of mutational events for four of the five tumors with time running down vertically from the root to the bifurcations of the phylogenetic tree. RCC-VTT-03 was excluded from the analysis with Canopy because of insufficient purity. Each sample consisted of a mixture of subclones that are represented as stacked bar plots (Fig. 4, right panels).

In three of the four tumors that entered the final analysis, there was a phylogenetic separation between primary tumor and VTT (Fig. 4, dendrograms on the left) suggesting an evolutionary distance between the two compartments. There were also noticeable differences in the number of subclones and subclonal composition modeled by Canopy. Whereas SNVs and CNAs in RCC-VTT-01 and -04 were modeled into three subclones with a relatively even admixture in primary tumor and VTT, RCC-VTT-02 and -05 showed a higher number of subclones (up to six) and greater differences in the subclonal composition of a given sample.



Figure 3. Mutational landscape of RCC with venous tumor thrombus. Functional SNVs (green) and indels (orange) in 483 genes are shown (grey indicates absence of a mutation). Numbers in cells describe the number of reads carrying the alternative allele and the respective sequencing coverage at this position. Bars on the right side of each plot indicate primary tumor-specific (red) or VTT-specific (blue) SNVs and indels. Driver mutations are highlighted in red.

We next analyzed whether the higher subclonal heterogeneity detected in RCC-VTT-02 and -05 may be associated with enhanced replication stress, a major cause of genomic instability and clonal heterogeneity³⁸. An immunohistochemical analysis of corresponding primary tumor and sections of the five sequenced patients for proliferative activity (Ki-67) and signs of an activated DNA damage response (γ H2AX; Suppl. Fig. 3) was performed. We found a consistently and significantly (p values between $p \leq 0.05$ and $p \leq 0.0005$) higher proliferation

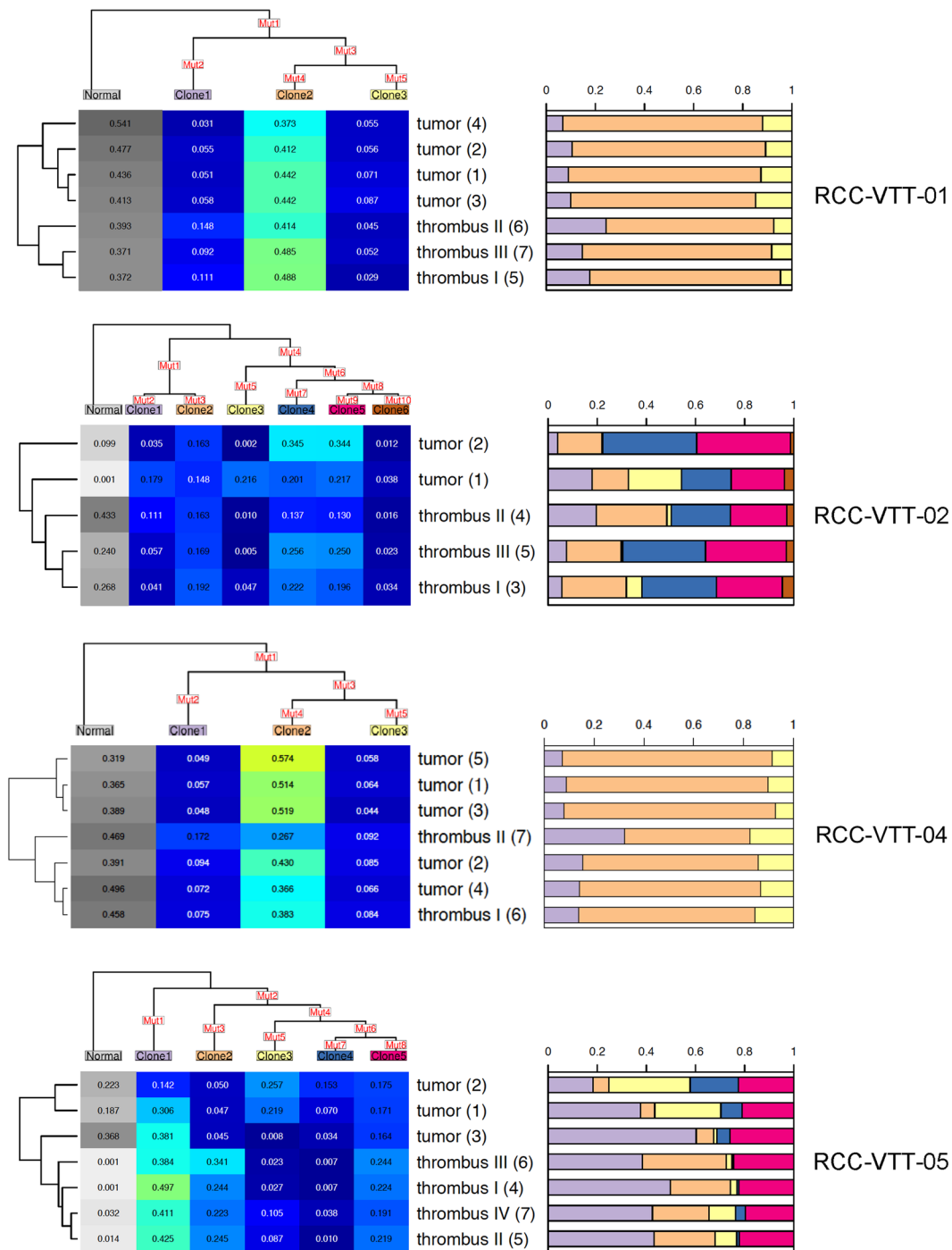


Figure 4. Clonal evolution in RCC with VTT. The clonal heterogeneity for all samples was determined using Canopy. Clonal evolution is depicted as a tree, the leaves of which correspond to the subclones denoted by “Clone1” up to “Clone6”. The number of leaves corresponds to the optimal number of subclones. The set of all mutations (SNVs and CNAs) among all samples of the patient is distributed along the tree in a (heuristically) optimal way (“Mut1”–“Mut10”). Each clone carries all the mutations, which can be collected when following the path from the respective leaf to the root of the tree. The matrix below the mutation tree shows the subclonal composition of each sample (values per line add up to 1). The leaf named “Normal” represents normal tissue contamination and has been neglected in further analyses. The contributions of each other subclone to a sample are illustrated in a stacked bar plot to the right of the matrix, the colors of which correspond to the colors of the leaf labels of the mutation tree. To the left of the matrix, the relationship between the samples with respect to subclonal composition (without “Normal”) is illustrated by a dendrogram.

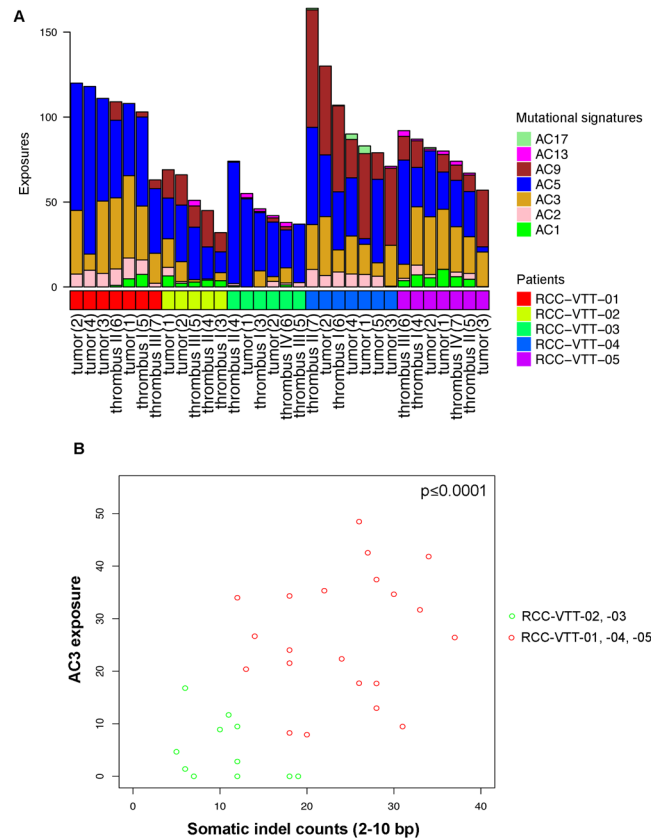


Figure 5. Mutational signatures in RCC with VTT. **(A)** Mutational signature analysis of the five patients. The shown signatures represent the following mutational processes: AC1 = spontaneous deamination; AC2 = APOBEC action; AC3 = homologous recombination repair defect; AC5 = unknown/metabolic stress; AC9 = Pol eta/somatic hypermutation; AC13 = APOBEC; AC17 = unknown. Note the presence of AC3 indicating defective homologous recombination repair in three of the five patients (RCC-VTT-01, -04, -05). **(B)** Scatterplot of exposures to signature AC3 and indel counts (2–10 bp) per sample. Wilcoxon rank-sum test between the two colour-coded groups, $p \leq 0.0001$.

rate in the venous thrombi in comparison to at least one of the respective primary tumor samples (Suppl. Fig. 3). At the same time, tumor cells with γ H2AX foci indicating an activated DNA damage response were detected in all primary RCCs and corresponding venous tumor thrombi at varying levels (Suppl. Fig. 3). Although a high number of γ H2AX foci-positive tumor cells were detected in RCC-VTT-05, which also showed a high clonal diversity, there were overall no fundamental differences between tumors with high clonal diversity in comparison to RCCs with lower clonal diversity. This finding suggests that replication stress may not play a major role in shaping clonal heterogeneity.

Taken together, these findings indicate that a phylogenetic separation of tumor and VTT can occur without new mutations suggesting that pre-existing subclones of the primary tumor drive VTT formation.

Mutational signatures in RCC with VTT. Mutational signatures are an emerging concept in understanding cancer genomes and underlying mechanisms of genomic instability²³. To further elucidate mutational processes in RCC with VTT, we performed a supervised analysis of mutational signatures using the software YAPSA²⁴, i.e. an analysis with pre-described mutational signatures taken from <http://cancer.sanger.ac.uk/cosmic/signatures> (Fig. 5A).

Mutational signatures can be grouped into different categories including clock-like signatures³⁹, signatures associated with mutagenic enzymatic systems or signatures associated with DNA damage repair defects²³. In the present cohort and in accordance with previous analyses of mutational signatures in RCC^{23,39}, we found a high prevalence of signature AC5 indicating metabolic stress in all samples (Fig. 5A). Moreover, we found exposure to signature AC3 indicating homologous recombination (HR) repair defects in multiple samples of three of the five patients (RCC-VTT-01, -04, -05; Fig. 5A). Signature AC3, which can be caused by mutations in DNA repair genes such as *BRCA1* or *BRCA2* but also others, has not been reported in RCC before^{23,39}, and may have important translational implications since it is known to confer “BRCAness”⁴⁰. Although no *BRCA1* or *BRCA2* mutations were found in the three RCCs with AC3 exposure, the presence of such mutations is not a prerequisite for “BRCAness”. Instead we detected mutations in *BAP1*, *CDK12*, *ERCC1* or *PTEN*, which have recently been implicated in “BRCAness”⁴⁰. In line with our notion of “BRCAness” in RCC, these genes were found mutated in tumors with mutational signature AC3 (RCC-VTT-01, 04, -05; Fig. 3).

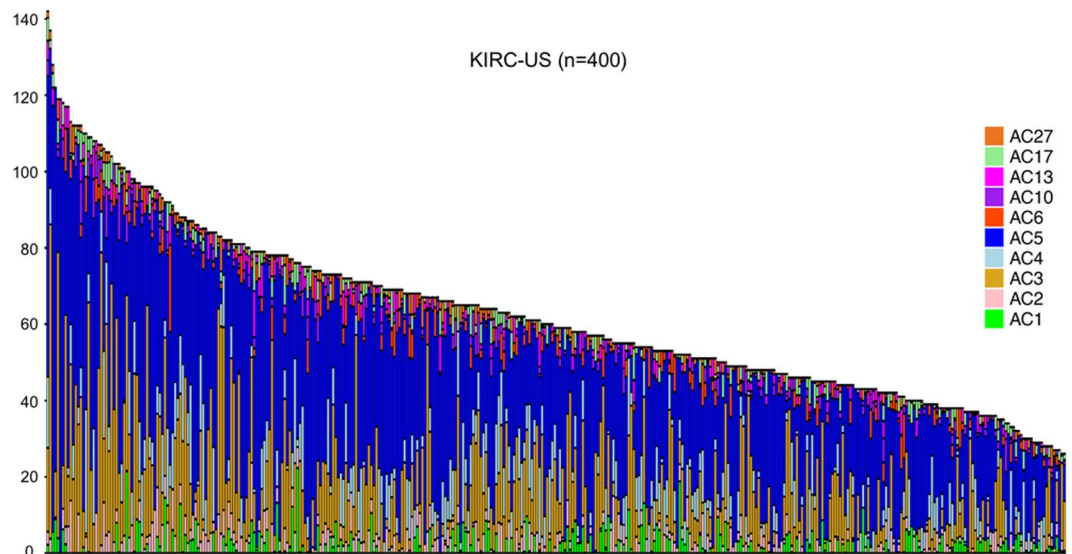


Figure 6. Mutational signatures in the TCGA KIRC-US cohort. Re-analysis of the TCGA KIRC-US cohort using 400 samples of the clear cell RCC TCGA cohort (KIRC-US; patients filtered by a minimum of 25 somatic SNVs per sample; <https://portal.gdc.cancer.gov/projects/TCGA-KIRC>). Note the high frequency of mutational signature AC3 and AC5. Each bar represents one patient. The y axis depicts the number of mutations. Additional signatures represent the following mutational processes: AC4 = smoking; AC6 = defective DNA mismatch repair; AC10 = altered POLE activity; AC27 = unknown.

To further corroborate the notion of “BRCAness” in tumors with signature AC3, we analyzed counts of small indels of 2–10 bp length. Small deletions of up to 50 bp have been reported to be characteristic for *BRCA1/2* mutation associated DNA repair defects^{41–43}. A significant correlation between exposure to mutational signature AC3 and higher somatic indel counts was detected (Wilcoxon rank-sum test $p \leq 0.0001$; Fig. 5B).

In addition, HRD (Homologous Recombination Deficiency) and LST (Large Scale Transition) scores were estimated using smoothed copy number profiles. These two scores, however, did not show a significant correlation with AC3 exposure.

Furthermore, we tested whether indel counts and AC3 exposure were dependent on the localization of the sample taken (i.e., primary tumor or VTT). No such correlation could be corroborated and we conclude that there is no modulating effect of the origin of the specimen in this respect.

Lastly, we sought to determine whether AC3 exposure is a general feature of clear cell RCC. To this end, we re-analyzed data from 400 patients of the clear cell RCC TCGA cohort (KIRC-US, patients filtered by a minimum of 25 somatic SNVs per sample; <https://portal.gdc.cancer.gov/projects/TCGA-KIRC>) and found AC3 exposure in 297 of 400 samples (74.25%; Fig. 6).

In summary, our results show that a subset of RCCs with VTT show signs of “BRCAness” including exposure to mutational signature AC3 as well as an increase of small indels. A re-analysis of mutational signatures in the TCGA KIRC-US cohort furthermore suggests that AC3 exposure is a frequent finding in clear cell RCC.

Discussion

Herein, we demonstrate that intravenous tumor thrombi in RCC are not simply an accumulation of tumor cells of uncertain viability and thrombotic material. Instead, our results indicate that they contain viable tumor cells with mutational heterogeneity between primary tumor and VTT. Moreover, we find signs for “BRCAness” in a subset of RCCs with VTT.

The mutational heterogeneity within individual RCCs, among patients and between primary tumors and venous thrombi is in line with previous reports^{1,2}. RCCs are also known to harbor more than one driver mutation, a finding that is confirmed in our study². Our results furthermore suggest that, despite mutational heterogeneity, no additional mutations are required for VTT formation and that pre-existing subclones of the primary tumor can drive this process.

“BRCAness” is a relatively recent concept in tumor biology that capitalizes on the fact that defects in error-free DNA double-strand break repair by HR repair represent a common outcome of inactivating mutations in a large number of genes in addition to *BRCA1* or *BRCA2*^{15,40}. The mutational signature AC3, which is associated with “BRCAness”, was not found in two large studies on mutational signatures that included RCCs^{23,39}. This most likely reflects technical differences since the algorithm to detect mutational signatures used in the present study is supervised and highly sensitive for detecting known mutational signatures. A re-analysis of the TCGA KIRC-US data confirmed a relatively high frequency of signature AC3 in clear cell RCC. Although we were not able to retrieve information on the VTT status in these tumors, the fact that over 70% of tumors showed AC3 exposure in comparison to the reported VTT frequency would argue that AC3 exposure is a more general finding in clear cell RCC and not specific for RCCs with VTT.

RCC is not known to frequently harbor mutations in DNA repair genes and is considered largely insensitive to DNA damaging agents. In the present study, we did not find any mutations in the “classical” HR repair genes such as *BRCA1* or *BRCA2* although such mutations do exist in RCC⁴⁴. However, we did find functional SNVs or indels in genes linked to either HR gene expression (*CDK12*) or execution/completion of HR (*PTEN*, *ERCC1*)⁴⁰. We hasten to add that other genes or combinations of genes that harbor functional SNVs and indels identified in our study could also provide a link to the “BRCAness” signature. Hence, a careful interpretation of the causative role of gene mutations is necessary and eventually requires a functional validation.

It is noteworthy that some genomic characteristics that have been found to be associated with “BRCAness”, namely the HRD and LST scores, did not show a correlation with mutational signature AC3 in our study. This finding could be related to the fact that whole exome sequencing was used, which may lead to an underestimation of these LOH-based scores or may reflect a less stringent correlation between signature AC3, “BRCAness” and LST/HRD scores in RCC in comparison to breast or ovarian cancer with germline *BRCA1/2* mutations.

“BRCAness” has been shown in other tumor entities to lead to an enhanced sensitivity to PARP inhibitors and platinum salts¹⁵. A key translational question that arises from our study is hence whether RCC patients with AC3 exposure and other signs of “BRCAness” will in fact benefit from such therapies.

Another mutational signature that was frequently detected in our samples is AC5. The unique characteristic of ccRCCs to generally downregulate important metabolic pathways and instead rely on a small metabolic network not usually found in other tumors has recently been reported⁴⁵. These metabolic changes were found to be associated with a loss of the *VHL* tumor suppressor on chromosome 3p. In four of five RCCs analyzed in our study, we found histologically a clear cell phenotype and a corresponding loss of chromosome 3p (Suppl. Fig. 2). All of these tumors showed a ubiquitous presence of mutational signature AC5. In only one tumor (RCC-VTT-04), AC5 was undetectable in one of the specimens (tumor (3)) and detected at a low level in another specimen (tumor (1)). This RCC was a papillary RCC type II with the characteristic *SETD2* and *NF2* mutation^{33,34}. The fact that AC5 was detectable in this tumor without *VHL* mutation suggests that other mechanisms may contribute to its origin as well. AC5 has been reported to be high in papillary RCC as well as in ccRCC³⁹ before and our results confirm this notion. AC5 has also been suggested to represent a clock-like signature³⁹. Patient RCC-VTT-04, the patient with papillary RCC type II and heterogeneous AC5 exposure, was the oldest of the five patients analyzed. This suggests that the tissue of origin and the cancer entity have an important modulating effect on and may even outweigh the clock-like nature of AC5. The biological basis of signature AC5, which can be detected in many human cancers, is still poorly understood and future studies are needed to fully understand the role of AC5 as well as other mutational signatures in tumor development and progression⁴⁶. Although it has recently been suggested that mutational signatures can have a certain degree of gender specificity⁴⁷, the study population used here is too small to draw any conclusions in that respect.

Our report on the mutational landscape of RCC with VTT using multiregion whole exome sequencing of the VTT provides important insights into the natural history of this unique growth pattern. Our results are in line with a previous report of heterogeneous mutations in a venous thrombus from a single RCC patient². We provide the evidence for a “BRCAness” signature in RCC, which may have important translational and clinical implications and may lead to novel treatment modalities such as PARP inhibitors, platinum salts and others in a selected group of patients. Pre-clinical studies to test this notion are urgently warranted.

References

- Gerlinger, M. *et al.* Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *N. Engl. J. Med.* **366**, 883–892 (2012).
- Gerlinger, M. *et al.* Genomic architecture and evolution of clear cell renal cell carcinomas defined by multiregion sequencing. *Nat. Genet.* **46**, 225–233 (2014).
- Pouliot, F., Shuch, B., Laroche, J. C., Pantuck, A. & Beldegrun, A. S. Contemporary management of renal tumors with venous tumor thrombus. *J. Urol.* **184**, 833–41; quiz 1235 (2010).
- Martínez-Salamanca, J. I. *et al.* Lessons learned from the International Renal Cell Carcinoma-Venous Thrombus Consortium (IRCC-VTC). *Curr Urol Rep* **15**, 404 (2014).
- Psutka, S. P. & Leibovich, B. C. Management of inferior vena cava tumor thrombus in locally advanced renal cell carcinoma. *Ther Adv Urol* **7**, 216–229 (2015).
- Bertini, R. *et al.* Impact of venous tumour thrombus consistency (solid vs friable) on cancer-specific survival in patients with renal cell carcinoma. *Eur. Urol.* **60**, 358–365 (2011).
- Weiss, V. L. *et al.* Prognostic significance of venous tumour thrombus consistency in patients with renal cell carcinoma (RCC). *BJU Int.* **113**, 209–217 (2014).
- Mager, R. *et al.* Renal cell carcinoma with inferior vena cava involvement: Prognostic effect of tumor thrombus consistency on cancer specific survival. *J Surg Oncol* <https://doi.org/10.1002/jso.24395> (2016).
- Reese, A. C., Whitson, J. M. & Meng, M. V. Natural history of untreated renal cell carcinoma with venous tumor thrombus. *Urol. Oncol.* **31**, 1305–1309 (2013).
- Kirkali, Z. & Van Poppel, H. A critical analysis of surgery for kidney cancer with vena cava invasion. *Eur. Urol.* **52**, 658–662 (2007).
- Haferkamp, A. *et al.* Renal cell carcinoma with tumor thrombus extension into the vena cava: prospective long-term followup. *J. Urol.* **177**, 1703–1708 (2007).
- Hoefflin, R. *et al.* Spatial niche formation but not malignant progression is a driving force for intratumoural heterogeneity. *Nat Commun* **7**, ncomms11845 (2016).
- Waclaw, B. *et al.* A spatial model predicts that dispersal and cell turnover limit intratumour heterogeneity. *Nature* **525**, 261–264 (2015).
- Almendo, V. *et al.* Genetic and phenotypic diversity in breast tumor metastases. *Cancer Res.* **74**, 1338–1348 (2014).
- Lord, C. J. & Ashworth, A. PARP inhibitors: Synthetic lethality in the clinic. *Science* **355**, 1152–1158 (2017).
- Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
- Jones, D. T. W. *et al.* Dissecting the genomic complexity underlying medulloblastoma. *Nature* **488**, 100–105 (2012).
- Jones, D. T. W. *et al.* Recurrent somatic alterations of FGFR1 and NTRK2 in pilocytic astrocytoma. *Nat. Genet.* **45**, 927–932 (2013).
- Wang, K., Li, M. & Hakonarson, H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* **38**, e164 (2010).

20. Sherry, S. T. *et al.* dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res.* **29**, 308–311 (2001).
21. 1000 Genomes Project Consortium *et al.* An integrated map of genetic variation from 1,092 human genomes. *Nature* **491**, 56–65 (2012).
22. Rimmer, A. *et al.* Integrating mapping-, assembly- and haplotype-based approaches for calling variants in clinical sequencing applications. *Nat. Genet.* **46**, 912–918 (2014).
23. Alexandrov, L. B. *et al.* Signatures of mutational processes in human cancer. *Nature* **500**, 415–421 (2013).
24. Hübschmann, D., Gu, Z. & Schlesner, M. YAPSA: Yet Another Package for Signature Analysis. R package version 1.0.0. (2015).
25. Wang, Y., Lawson, C. L. & Hanson, R. J. *Isei*: Solving Least Squares Problems under Equality/Inequality Constraints. *R package version 1*, 1–1 (2015).
26. Gehringer, J. S., Fischer, B., Lawrence, M. & Huber, W. SomaticSignatures: inferring mutational signatures from single-nucleotide variants. *Bioinformatics* **31**, 3673–3675 (2015).
27. Talevich, E., Shain, A. H., Botton, T. & Bastian, B. C. CNVkit: Genome-Wide Copy Number Detection and Visualization from Targeted DNA Sequencing. *PLoS Comput. Biol.* **12**, e1004873 (2016).
28. Olshen, A. B. *et al.* Parent-specific copy number in paired tumor-normal studies using circular binary segmentation. *Bioinformatics* **27**, 2038–2046 (2011).
29. Popova, T. *et al.* Ploidy and large-scale genomic instability consistently identify basal-like breast carcinomas with BRCA1/2 inactivation. *Cancer Res.* **72**, 5454–5462 (2012).
30. Abkevich, V. *et al.* Patterns of genomic loss of heterozygosity predict homologous recombination repair defects in epithelial ovarian cancer. *Br. J. Cancer* **107**, 1776–1782 (2012).
31. Jiang, Y., Qiu, Y., Minn, A. J. & Zhang, N. R. Assessing intratumor heterogeneity and tracking longitudinal and spatial clonal evolutionary history by next-generation sequencing. *Proceedings of the National Academy of Sciences* **113**, E5528–37 (2016).
32. Cancer Genome Atlas Research Network. Comprehensive molecular characterization of clear cell renal cell carcinoma. *Nature* **499**, 43–49 (2013).
33. Cancer Genome Atlas Research Network *et al.* Comprehensive Molecular Characterization of Papillary Renal-Cell Carcinoma. *N. Engl. J. Med.* **374**, 135–145 (2016).
34. Durinck, S. *et al.* Spectrum of diverse genomic alterations define non-clear cell renal carcinoma subtypes. *Nat. Genet.* **47**, 13–21 (2015).
35. Zhang, D., Zaugg, K., Mak, T. W. & Elledge, S. J. A role for the deubiquitinating enzyme USP28 in control of the DNA-damage response. *Cell* **126**, 529–542 (2006).
36. Niedernhofer, L. J. *et al.* The structure-specific endonuclease Ercc1-Xpf is required to resolve DNA interstrand cross-link-induced double-strand breaks. *Mol. Cell. Biol.* **24**, 5776–5787 (2004).
37. Song, W., Pascal, J. M., Ellenberger, T. & Tomkinson, A. E. The DNA binding domain of human DNA ligase I interacts with both nicked DNA and the DNA sliding clamps, PCNA and hRad9-hRad1-hHus1. *DNA Repair (Amst.)* **8**, 912–919 (2009).
38. Burrell, R. A. *et al.* Replication stress links structural and numerical cancer chromosomal instability. *Nature* **494**, 492–496 (2013).
39. Alexandrov, L. B. *et al.* Clock-like mutational processes in human somatic cells. *Nat. Genet.* **47**, 1402–1407 (2015).
40. Lord, C. J. & Ashworth, A. BRCAness revisited. *Nat. Rev. Cancer* **16**, 110–120 (2016).
41. Tutt, A. *et al.* Mutation in Brca2 stimulates error-prone homology-directed repair of DNA double-strand breaks occurring between repeated sequences. *EMBO J.* **20**, 4704–4716 (2001).
42. Xia, F. *et al.* Deficiency of human BRCA2 leads to impaired homologous recombination but maintains normal nonhomologous end joining. *Proc. Natl. Acad. Sci. USA* **98**, 8644–8649 (2001).
43. Moynahan, M. E., Pierce, A. J. & Jasin, M. BRCA2 is required for homology-directed repair of chromosomal breaks. *Mol. Cell* **7**, 263–272 (2001).
44. Dietz, S. *et al.* Patient-specific molecular alterations are associated with metastatic clear cell renal cell cancer progressing under tyrosine kinase inhibitor therapy. *Oncotarget* <https://doi.org/10.18632/oncotarget.18200> (2017).
45. Gatto, F., Nookaew, I. & Nielsen, J. Chromosome 3p loss of heterozygosity is associated with a unique metabolic network in clear cell renal carcinoma. *Proceedings of the National Academy of Sciences* **111**, E866–75 (2014).
46. Supek, F. & Lehner, B. Clustered Mutation Signatures Reveal that Error-Prone DNA Repair Targets Mutations to Active Genes. *Cell* **170**, 534–547.e23 (2017).
47. Letouzé, E. *et al.* Mutational signatures reveal the dynamic interplay of risk factors and cellular processes during liver tumorigenesis. *Nat Commun* **8**, 1315 (2017).
48. Novick, A. C., Stroom, S. B., Pontes, J. E. & Stewart, B. H. *Stewart's Operative urology.* (1989).

Acknowledgements

We would like to thank the tissue bank of the National Center for Tumor Diseases Heidelberg for the procurement of tissue samples. The results shown here are in part based upon data generated by the TCGA Research network: <http://cancergenome.nih.gov/>. This work was supported by the Wilhelm Sander-Stiftung, the Medical Faculty Heidelberg and an internship by the International Federation of Medical Students Associations (to L.V.C.). D.H. is a member of the Hartmut Hoffmann-Berling International Graduate School of Molecular and Cellular Biology (HBIGS) and of the M.D./Ph.D.-program of the University of Heidelberg.

Author Contributions

S.D., M.H., H.S. and M.S. conceived and jointly supervised the study; G.W., D.H. and K.K. performed the bioinformatic analyses; D.H. performed the analysis of mutational signatures; Y.T., M.He., L.V.C., L.T. performed experiments; C.N., C.P., C.Ga., T.K., J.N.D., G.H., D.T., B.H., S.Pa., D.J., and C.Gr. enrolled subjects and collected and contributed clinical data; K.W., A.S., and W.R. contributed pathology information; S.P. and A.D. gave important advice on data analysis and interpretation; R.E. and M.S. supervised the bioinformatics analyses; S.D. supervised the research and wrote, together with G.W. and D.H., the first draft of the manuscript; all authors contributed to the editing of the manuscript and approved the final version.

Additional Information

Supplementary information accompanies this paper at <https://doi.org/10.1038/s41598-018-25544-z>.

Competing Interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018