**METHODS**

**Patients and sample procurement**

Tumor tissue and matched-blood from six kidney cancer patients who underwent partial or radical nephrectomy at our institution were obtained as follows. Peripheral venous blood samples were drawn in the clinic during pre-surgical visits and used as matched germline controls. Primary tumor specimens were procured during the surgical procedure and six *ex-vivo* core-needle biopsy samples were obtained from spatially-separated regions. All samples were snap-frozen and stored until genomic material extraction. Specimens were also reviewed by a genitourinary pathologist (Y.C.) and a diagnosis of clear cell renal cell carcinoma was established in all cases. Median tumor size (maximum diameter, in cm, on pathologic review) was 6.75 cm (range, 4 to 9.5) and five of the tumors were pT3 stage (**Supplementary table 1**). One of the tumors (RCC006) did not show any non-silent mutations in any of the regions and was excluded from further analysis. The final sample size consisted of 30 tumor regions from five patients as well as five tumor DNA pools (each one containing DNA from six regions).

**Ethics**

The institutional ethics review board pre-approved all research-related activities. All individuals provided written informed consent for the participation in this study, which included their tumor and normal tissue being profiled and their clinical data being shared in a non-identifiable manner. All activities related to this study were conducted in accordance with the principles of Good Clinical Practice and the Declaration of Helsinki.

**DNA extraction and multiregional pool construction**

Samples from each tumor region were stored at -80\*C for 3-4 weeks. After thawing the samples in regular ice, DNA extraction was performed using the DNeasy Blood & Tissue Kit according to the manufacturer’s protocol (QIAGEN catalog #69504). Genomic DNA from tumor regions and blood samples was eluted in 50 ml of nuclease-free water. Tumor DNA pools were confected by combining 10 ml of the eluate from each region and mixing them by pipetting into a single 60 ml sample. A total of 8 DNA samples from each patient (1 blood, 6 tumor regions and 1 tumor pool) were sent for sequencing at the Integrated Genomics Operations Core of Memorial Sloan Kettering Cancer Center, a CLIA-certified laboratory (see Acknowledgments).

**Sequencing, raw data processing and somatic mutation calling**

Sequencing was done using MSK-IMPACT®, a targeted sequencing panel approved by the FDA for the study of solid tumors. It captures the exons of 341 cancer genes as well as a set of known single-nucleotide polymorphisms (SNPs) tiled across the genome (for CN estimation)[[4]](https://paperpile.com/c/cgubWB/wMvd). The median sequencing depth for tumor and normal samples was 876x (range, 574x - 1911x) and 544x (range, 500x - 802x), respectively.Raw sequencing data in FASTQ format was aligned to the human genome (version b37) using the Burrows–Wheeler aligner (BWA v.0.7.10). Local realignment was performed with the Genome Analysis Toolkit (GATK v4.1.4.1[[17]](https://paperpile.com/c/cgubWB/1FdV2)) and duplicate reads removed using Picard v.2.13.

For somatic variant calling, we used our integrated bioinformatics pipeline which includes four different variant calling tools: MuTect2 (part of GATKv4.1.4.1[[17]](https://paperpile.com/c/cgubWB/1FdV2)), Strelka2 v2.9.10[[18]](https://paperpile.com/c/cgubWB/AQlkG), Varscan v2.4.3 [[19]](https://paperpile.com/c/cgubWB/hW82n) and Platypus [[20]](https://paperpile.com/c/cgubWB/bUFVf). Ancillary filters were used to obtain high-accuracy calls, these included: coverage of at least 10x in the tumor, with 5 or more reads supporting the variant, a variant allele frequency (vAF, i.e. proportion of DNA read ) greater or equal than 5% in single-region and 2% in pooled tumor samples, and a vAF below 2% in the matched normal sample. Only nonsynonymous somatic mutations were considered, and single-nucleotide variants (SNVs) identified at a frequency higher than 1% in dbSNP[[21]](https://paperpile.com/c/cgubWB/dzPcP) or 1000Genomes project[[22]](https://paperpile.com/c/cgubWB/VcL0A) were excluded. Finally, dubious calls were manually reviewed by investigators using the Integrative Genome Viewer software v2.4.10 for additional accuracy [[23,24]](https://paperpile.com/c/cgubWB/OKhhN+RBQ8O). A total of 23 unique nonsynonymous somatic mutation events were considered in the analysis. Tumor mutational burden (TMB) was defined as the total number of mutations identified in a tumor specimen (considering all available regions).

**Allele-specific copy-number (ASCN) analysis**

Allele-specific copy number analysis (ASCN) was performed using FACETS v0.5.6 [[25]](https://paperpile.com/c/cgubWB/ZYhWa) through a publicly-available R package (<https://github.com/mskcc/facets/>). By comparing the tumor and copy-number (CN) profiles, this tool provides integer copy number (CN) values for each tumor allele in addition to other genome-wide estimates, such as tumor ploidy and sample purity (i.e. the proportion of DNA in tumor samples derived from tumor cells).

**Clonality assessment: cancer-cell fraction (CCF) estimation**

We used the cancer-cell fraction (CCF) of each somatic variant as a surrogate of the relative timing of each mutation. This measure represents the proportion of tumor cells bearing a somatic variant of interest, and has been extensively used as a clonality surrogate in several landmark cancer genomic studies[[26–28]](https://paperpile.com/c/cgubWB/gu00+VwWX+wTts). The CCF of each variant was calculated assuming a diploid CN state (i.e. total CN=2, minor allele CN=1) in the normal tissue (CPNnorm) and using the vAF, locus-specific ploidy (CPNmut) and sample purity (p) information as described by McGranahan et al. [[29]](https://paperpile.com/c/cgubWB/I68Z):

**Cell abundance estimates in pooled DNA samples**

We derived the formulas to estimate sample purity and mutation CCF values in a hypothetical tumor DNA pool using multi-regional data (**Supplementary Material 1b**). Although the expected purity of a tumor pool is a near-average of the value in each region, CCF cannot be estimated the same way. Owing to the fact that total DNA from different tumor samples was combined, and that CCF represents a measure of mutation abundance relative to the cancer cell compartment (rather than all cells), single-region CCF averages are not appropriate to estimate the values in the pool. Therefore, we considered the mutant cell fraction (MCF, i.e. proportion of mutant cells, out of all cells) and purity of each region as follows:

**TRACERx analyses**

We obtained publicly-available somatic mutation data from the original RCC and NSCLC studies of the TRACERx consortium. Individuals who met the following criteria were included in the analysis: i) absence of bilateral or multifocal renal tumors, and ii) more than one primary tumor region with non-silent mutations available. The final RCC cohort included 92 individuals (1146 regions, profiled with a sequencing panel targeting 110 cancer genes), while the NSCLC cohort included 95 individuals (292 regions) profiled using exome sequencing **(Supplementary figure 3a)**. Mutation-annotation files (MAF) containing the CCF and the cell-abundance estimates used in the analysis were provided (**Supplementary table 4, 5**).

**Somatic variant annotation**

Nonsynonymous exonic somatic mutation data both from the in-house samples and TRACERx cohorts were parsed into MAF format using the maf2maf function (v2.4) from a publicly-available package (<https://github.com/mskcc/vcf2maf>). The MAFs were then annotated with the corresponding CCF data using an R implementation of the previously-described formula (<https://github.com/mskcc/facets-suite>). Finally, the evidence on the treatment implications of each mutation was assessed using OncoKB [[8]](https://paperpile.com/c/cgubWB/YCkj) through its Python implementation (<https://github.com/mskcc/facets-suite>). Only levels of evidence 1 through 3 were included in the analysis, and the database was queried on March 6th of 2020.

**Clonality definitions**

Clonality definitions in sequencing studies vary widely and different methods and definitions have been used in single and multi-regional sequencing [[9]](https://paperpile.com/c/cgubWB/M6vD). Most commonly, specific thresholds are set in the CCF continuum to classify mutations into ‘clonal’ or ‘subclonal’, with some investigators opting for a third ‘indeterminate’ category. We opted for a binary classification approach using the same clonality definitions from the original TRACERx Renal publications (i.e. mutations had to be present in every region at a CCF greater or equal to 50% to be considered clonal) [[6,30]](https://paperpile.com/c/cgubWB/6ai7+tOiA). Similarly, a threshold of CCF≥75% was used to define clonality in the tumor DNA pools **(Supplementary Figure 2d and 4b)**.

Only variants at or above 5% CCF were reported in the TRACERx studies and therefore included in the analyses. To simulate subclonal mutation dropout due to excessive dilution of low frequency variants in DNA pools, we set a detection limit threshold in the simulated pools. Variants predicted to have a CCF below 2% in the tumor DNA pools were considered ‘missed’ in the analysis and counted as dropped (not considered in clonality analysis).

**Bootstrapping procedure**

To compare the outcomes of conventional and pooled sequencing, a bootstrapping procedure with 100 iterations was performed using nonsynonymous somatic mutation data from the TRACERx consortium. Analyses were conducted separately in the RCC and NSCLC cohorts due to differences in the breadth of sequencing. Briefly, given a number of tumor regions *R*, we obtained a random sample (with replacement) containing 70% of the tumors with at least *R* number of regions profiled (**Supplementary Figure 3a)**. We then created a ground-truth set of variant calls from all the regions available for these specimens and, mimicking random tissue sampling, selected a random subset of *R* regions to analyze. The variants detected and the clonality assertions made using only the chosen regions were compared to the previously-established ground truth (*i.e.* using all available regions); this was done both conventionally (i.e. multiple separate regions assessed) and simulating a tumor DNA pool (i.e. containing DNA from *R* regions mixed in equal proportions). Next, a series of variant- and tumor-level outcomes are calculated in each random subset of tumors, and estimates from the 100 iterations are averaged. This approach was repeated with an increasing number of regions and the estimates compared between conventional and pooled sequencing using parametric statistical testing.

**Outcomes and measures**

*Mutational dropout and misattribution of clonality*

We used a combination of variant- and tumor-level measures to compare the conventional and pooled sequencing approaches. Results are shown with an increasing number of regions in both real and simulated DNA pools **(Figure 1, 2)** as well as a comparable multiregional assessment of the same number of regions **(Supplementary Figure 3)**. In all the analyses, a ground truth set of variant calls was created using all the tumor regions available. True clonal mutations were defined as those with an estimated CCF≥0.5 in every region of the tumor. After listing all the mutations in a tumor and their true clonal status, we compared the calls to the results obtained from each sequenced sample (*in situ*) or a set of regions selected at random from the original data (*in silico*). Only mutations detected by the sample(s) of interest were considered when evaluating the clonality assertions made with both approaches.

The proportion of variants dropped and misclassified in each sample sequenced in-house was calculated, the sample-specific estimates were then averaged and compared between the individual regions and pools. For the *in silico* analysis, after obtaining the ground-truth variant calls from all the tumors in a given iteration of the bootstrapping, we calculated three measures for each of the outcomes: i) the absolute proportion of variants dropped/misclassified in all the pools selected (event-level), ii) the proportion of tumors with at least one variant dropped/misclassified (tumor-level) and, iii) the average proportion of variants dropped/misclassified per tumor (tumor-level average) (Supplementary Figure 3).

*Performance evaluation of pooled sequencing.*

Next, we evaluated the performance of the CCF values estimated from pooled samples as a binary classifier of mutation clonality. Each mutation was assigned a true ‘clonal’ or ‘subclonal’ status, and these assertions were contrasted to the CCF estimates obtained from the pooled samples using the aforementioned definitions. We calculated three different estimates (at the event-level) to assess the reliability of CCF in a pooled sequencing context; results are shown as average estimates (and their 95% CIs) across the 100 iterations of the bootstrapping. To evaluate the performance of a pooled CCF value without setting a specific threshold, area under the curve (AUC) estimates were calculated for the Receiver-Operating Characteristic (ROC) and Precision-Recall (PR) curves, using the R package ‘PRROC’. Given the frequent presence of unbalanced confusion matrices in this context (only branched or only truncal mutations present), we selected the Matthew’s Correlation Coefficient (MCC) to evaluate the performance of pooled CCF estimates as binary clonality classifiers [[31]](https://paperpile.com/c/cgubWB/4pxJ). The MCC estimates are shown across all possible CCF thresholds, and results were contrasted to the definitions used in the study (i.e. CCF≥0.5 in each region for ground-truth and CCF≥0.75 in pooled samples).

*Cost-effectiveness*

The cost-effectiveness of pooled sequencing and *bona fide* multiregional sequencing were compared as follows **(Figure 2b).** First, the total costs of profiling were calculated with both approaches using estimates from our own institution **(Supplementary Figure 5a)**. Then, the average dropout and clonality misattribution estimates were calculated across the 100 simulated samples **(Supplementary Figure 5b, c)**. Cost-effectiveness was defined as the average proportion of variants detected/correctly classified per tumor divided by the total cost of profiling. Finally, results were expressed relative to single-region profiling (relative cost-effectiveness).

*Risk misattribution rate*

To assess the translational value of the proposed approach, we evaluated the ability of pooled sequencing to correctly attribute risk points to each tumor, using the genomic criteria described by Voss et al. [[7]](https://paperpile.com/c/cgubWB/jf8d). A total of three genes were evaluated to risk-stratify patients:

1. Presence of a *BAP1* or *TP53* mutation - 1 point
2. Absence of *PBRM1* variants, or co-occurrence with one of the above - 1 point

Each tumor was assigned a ‘true risk score’ (of maximum two points) using all the regions available, and these were contrasted to the assertions made with the regions of interest. The rate of risk misattribution was defined as the proportion of tumors in the sample with erroneous risk scores (regardless of direction).

*Molecular subtype assignment (according to TRACERx Renal)*

We evaluated the ability of pooled sequencing to correctly assign a molecular subtype to each tumor as described by Turajlic et al. [[6]](https://paperpile.com/c/cgubWB/6ai7). Since one of the subtypes is characterized by presence of copy-number alterations, which were not evaluated as part of this study (PBRM1-->sCNA), minor adjustments had to be made in the way tumors were classified. Therefore, we could only achieve a concordance of 96% with the original results. First, a list of core driver genes (*VHL, PBRM1, SETD2, BAP1 and PTEN)* and PI3k pathway genes (*PIK3CA, MTOR, PTEN, TSC1, TSC2*) were obtained from the original study. Next, a set of rules were applied in hierarchical order to each tumor based on the mutations detected and their clonality estimates. If none of the rules were met, tumors were assigned to the ‘Non-driver’ (or indeterminate) sub-group. The rules were applied as follows:

i. ***Multiple clonal driver:*** presence of ≥ 2 *BAP1, PBRM1, SETD2* or *PTEN* clonal mutational events. *If not, then:*

ii. ***BAP1-driven:*** presence of a *BAP1* mutational driver event, and no other “core” mutational drivers in the same clone/subclone (other than *VHL*). In pooled samples, tumors were assigned to this category if they had a clonal *BAP1* variant or a subclonal one, with the highest CCF among all the drivers present (if any other were present). *If not, then:*

iii. ***PBRM1***→***SETD2:*** presence of a *PBRM1* mutation followed by a *SETD2* one. For tumor pools this meant a higher estimated CCF for the *PBRM1* variant compared to the one on *SETD2*. *If not, then:*

iv. ***PBRM1***→***PI3k:*** presence of a *PBRM1* mutation followed by a PI3k pathway variant. For tumor pools this meant a higher estimated CCF for the *PBRM1* variant compared to the one on the PI3k pathway gene(s). *If not, then:*

v. ***PBRM1***→**CNA:** presence of a *PBRM1* mutation followed by a driver somatic CNA. Since the CNAs were not evaluated, tumors were put tentatively in this category when a *PBRM1* variant was observed without any other mutational drivers (except for *VHL*). However, if the tumor met the criteria for the ‘*VHL* wild-type’ subtype, they were assigned that group instead. *If not, then:*

vi. ***VHL wild-type:*** absence of *VHL* mutation. *If not, then:*

vii. ***VHL monodriver:*** *VHL* as the only “core” driver mutation.

*TMB error (underestimation) rate*

TMB estimates were only reported for the NSCLC TRACERx cohort due to the low number of genes assessed in the RCC cohort (~110), which precludes accurate TMB assessment, as well as the lack of clinical value in this setting [[32]](https://paperpile.com/c/cgubWB/cBcA). Differences in TMB detection were compared between conventional and pooled sequencing by calculating an average TMB error estimate across the 100 simulated cohorts. For each patient, this was calculated by subtracting the TMB from a single region or a pool (of >1 region) to the ‘true TMB’ (i.e. the sum of all unique events), and the result expressed relative to the ground truth (i.e. true - observed/true).

**Statistical analyses**

Outcomes were compared between different sequencing approaches using parametric statistical testing. Average estimates were compared between conventional and pooled sequencing using *t*-tests. Unpaired tests were used given the different samples or random sets of patients used in the *in-situ* and *in-silico* analyses, respectively. Pearson correlation tests were used to compare observed versus expected cell-abundance estimates in the pooled samples sequenced in this study. Sensitivity analyses were performed to assess the performance of pooled CCF estimates for attributing clonality to mutations across different conditions. First, we explored the performance of CCF estimates of tumor pools to classify clonality using increasing tumor regions **(Supplementary figure 4a, b)**. Next, we tested the effect of changing the definitions of true clonality. This was done by setting increasing CCF thresholds to define ‘true clonality’ in the conventional multi-regional assessment used as reference (i.e. CCF≥10%, 50% and 90%, respectively, **Supplementary Figure 4c, d**). Results were reported as point estimates along with their 95% confidence intervals (i.e. averages across 100 simulations). Hypothesis tests were always two-sided and statistical significance defined as a p-value below 0.05. All analyses were performed using the R platform v3.5.3.