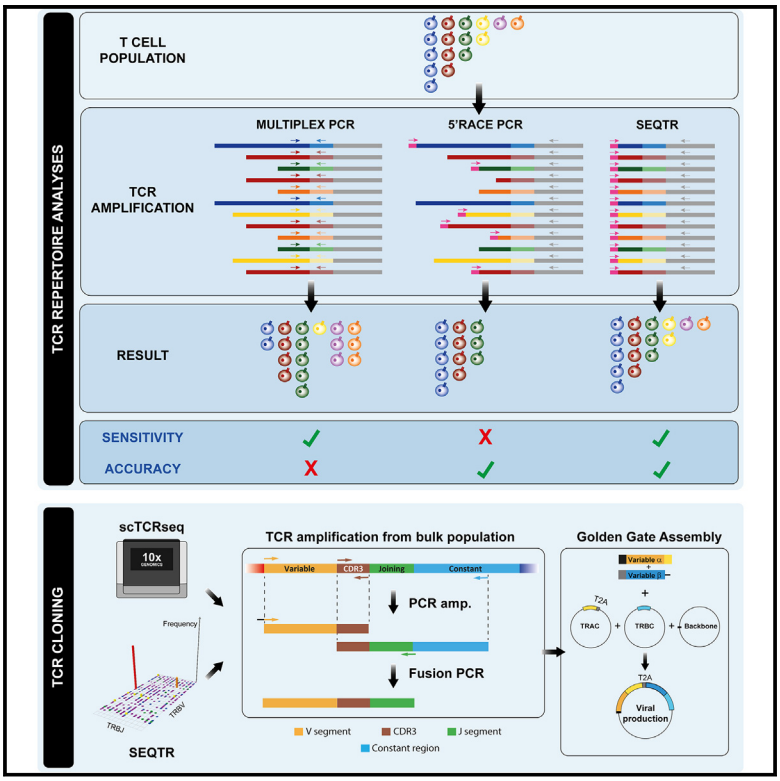
Cell Reports
Methods

Article

TCR sequencing and cloning methods for repertoire analysis and isolation of tumor-reactive TCRs

Graphical abstract



Highlights

- TCR RNA expression doesn't affect clonotype quantification when analyzing TCR repertoire
- SEQTR provides a sensitive and quantitative assay for TCR repertoire analyses
- Different TCR repertoire assays significantly affect repertoire metrics
- Direct TCR amplification from bulk population is time- and cost-effective for TCR cloning

Authors

Raphael Genolet, Sara Bobisse, Johanna Chiffelle, ..., Daniel E. Speiser, George Coukos, Alexandre Harari

Correspondence

raphael.genolet@chuv.ch (R.G.),
george.coukos@chuv.ch (G.C.),
alexandre.harari@chuv.ch (A.H.)

In brief

Current tools for T cell receptor (TCR) repertoire analyses remain expensive and/or potentially inaccurate. Genolet et al. present SEQTR, a method combining *in vitro* transcription and single primer pair TCR amplification for sensitive and quantitative repertoire analysis. Combined with simple PCR to amplify TCR from bulk population, it allows efficient tumor-specific TCR identification and cloning.

Genolet et al., 2023, Cell Reports Methods *3*, 100459
April 24, 2023
https://doi.org/10.1016/j.crmeth.2023.100459

CellPress

Cell Reports Methods

CellPress
OPEN ACCESS

Article

TCR sequencing and cloning methods for repertoire analysis and isolation of tumor-reactive TCRs

Raphael Genolet,[1,2,4,*] Sara Bobisse,[1,2] Johanna Chiffelle,[1,2] Marion Arnaud,[1,2] Rémy Petremand,[1,2] Lise Queiroz,[1,2] Alexandra Michel,[1,2] Patrick Reichenbach,[1] Julien Cesbron,[1,2] Aymeric Auger,[1,2] Petra Baumgaertner,[1,2] Philippe Guillaume,[1,2] Julien Schmidt,[1,2] Melita Irving,[1] Lana E. Kandalaft,[1,2] Daniel E. Speiser,[1] George Coukos,[1,3,*] and Alexandre Harari[1,3,*]

[1]Ludwig Institute for Cancer Research, Lausanne Branch, University of Lausanne and Lausanne University Hospital, Lausanne, Switzerland
[2]Center of Experimental Therapeutics, Department of Oncology, Lausanne University Hospital, Lausanne, Switzerland
[3]These authors contributed equally
[4]Lead contact
*Correspondence: raphael.genolet@chuv.ch (R.G.), george.coukos@chuv.ch (G.C.), alexandre.harari@chuv.ch (A.H.)
https://doi.org/10.1016/j.crmeth.2023.100459

MOTIVATION T cell receptor (TCR) analyses are instrumental to profile T cell repertoires and dynamics in many clinical settings. However, the time, cost, or technical biases of current technologies impair data accuracy and limit their inputs for the understanding of cellular immune responses and efficient development of novel therapies. Here, we present SEQTR, a TCR repertoire assay with improved sensitivity and accuracy relative to currently available methods. Furthermore, any specific TCR can be amplified and cloned from a bulk population in a time- and cost-effective way.

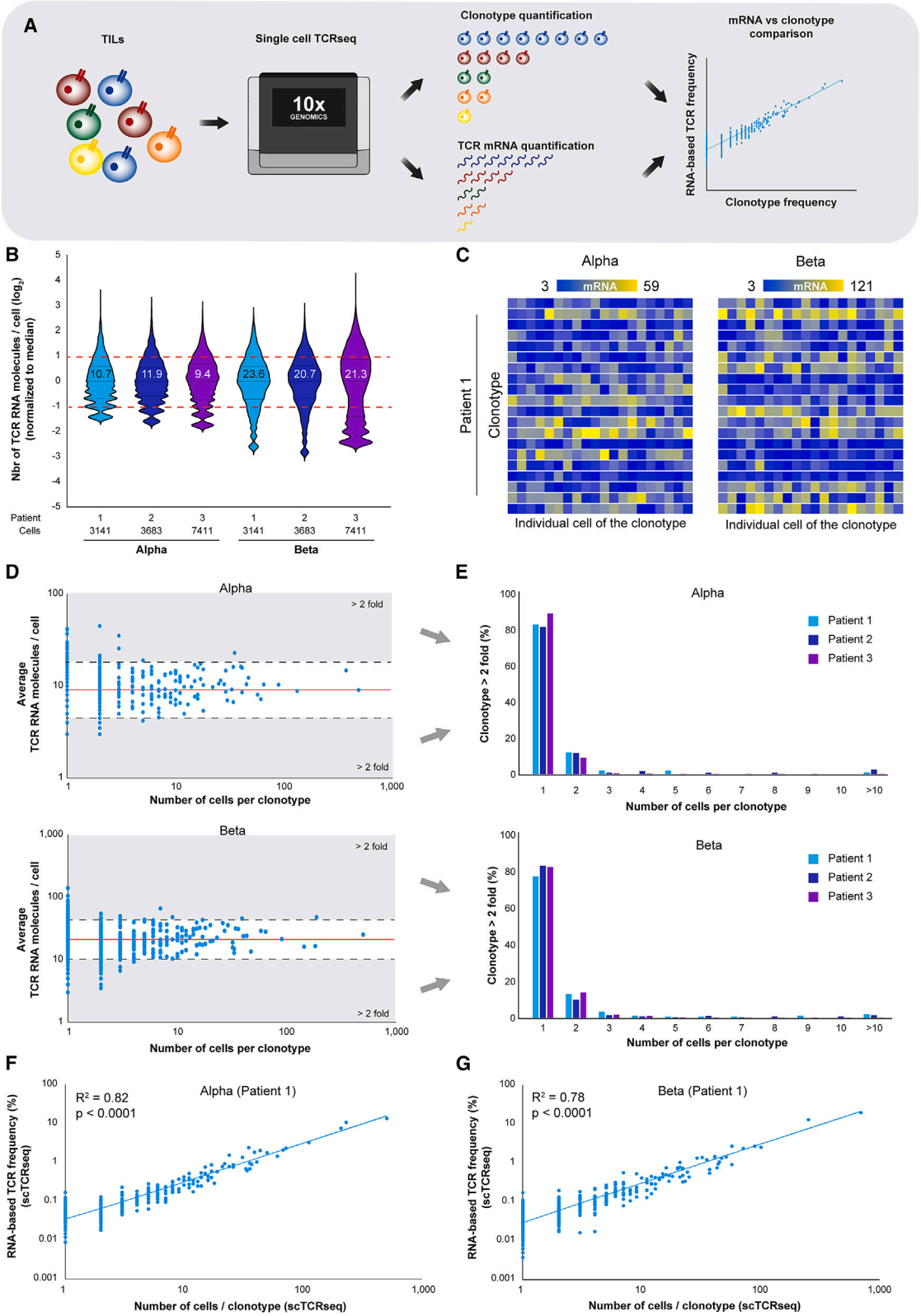
SUMMARY

T cell receptor (TCR) technologies, including repertoire analyses and T cell engineering, are increasingly important in the clinical management of cellular immunity in cancer, transplantation, and other immune diseases. However, sensitive and reliable methods for repertoire analyses and TCR cloning are still lacking. Here, we report on SEQTR, a high-throughput approach to analyze human and mouse repertoires that is more sensitive, reproducible, and accurate as compared with commonly used assays, and thus more reliably captures the complexity of blood and tumor TCR repertoires. We also present a TCR cloning strategy to specifically amplify TCRs from T cell populations. Positioned downstream of single-cell or bulk TCR sequencing, it allows time- and cost-effective discovery, cloning, screening, and engineering of tumor-specific TCRs. Together, these methods will accelerate TCR repertoire analyses in discovery, translational, and clinical settings and permit fast TCR engineering for cellular therapies.

INTRODUCTION

T cell receptor (TCR) repertoire analyses are extensively used to monitor T cell dynamics in many clinical settings of transplantation, autoimmune diseases, infection, and cancer.[1–4] T cell monitoring is important in following disease progression, assessing responses to treatment, or stratifying patients. Furthermore, in the field of cancer, recognition of tumor antigens by endogenous TCRs is associated with clinical benefit, while adoptive transfer of TCR-engineered cells has emerged as a promising therapeutic approach,[5–7] generating great interest for the discovery of tumor antigens and their cognate TCRs. For all these applications, timelines, cost, or biases introduced by current TCR technologies impair data accuracy and limit their inputs for the understanding of cellular immune responses and the efficient development of novel therapies.[1,8]

The TCR α/β heterodimers expressed on most T cells result from gene rearrangement of the variable (V), diversity (D), and joining (J) segments.[9] While diversity is a fundamental intrinsic feature of an efficient immune system, it remains challenging to capture it with common TCR repertoire tools. Next-generation sequencing (NGS) has considerably improved the ability to tackle TCR repertoire diversity, and NGS-based methods enabled by commercially available kits or services, such as multiplex PCR and 5′-RACE (rapid amplification of cDNA ends), are now commonly used.[10–13] However, multiplex PCR is known to introduce amplification bias due to differential primer efficiency,[14,15] and 5′-RACE is associated with poor efficiency of the template switch that adds the 5′ adapter in only 20%–60% of RNA molecules.[16] This may lead to less accurate quantification and a reduced sensitivity with precarious detection of low-frequency TCRs. Therefore, recent efforts have been made to



(legend on next page)

improve the accuracy and sensitivity of TCR repertoire methods.[17–19] The development of single-cell TCR sequencing (scTCR-seq) has opened opportunities to obtain more precise information on individual clones including TCR identification, yet the number of cells analyzed is limited such that single-cell technologies are not appropriate to reach the deepness of analyses required for some clinical monitoring. Similarly, cloning and screening strategies remain mandatory steps to validate specificities of TCRs identified with scTCR-seq. Here, we describe: (1) SEQTR (SEQuencing T cell Receptor), a quantitative and sensitive method to analyze human and murine T cell repertoires that circumvents the aforementioned biases; and (2) a cloning strategy to reduce timelines and resources needed to isolate TCR sequences of interest. Used separately or in combination, we illustrate how SEQTR could improve patient stratification using TCR metrics and describe a cost- and time-effective pipeline to identify, clone, and validate tumor-specific TCRs in only a few days.

RESULTS

RNA-based assays allow clonotype quantification

Both DNA-based and RNA-based TCR-seq assays have pros and cons. DNA is more stable and present at fixed copy numbers per cell, facilitating clonotype quantification. However, given the presence of DNA from irrelevant V and J segments that are not part of the rearranged sequences,[20] DNA-based assays may have decreased signal-to-noise ratio and reduced amplification efficiency. Moreover, DNA includes both TCRβ alleles, whereas most T cells express only one allele following allelic exclusion.[21] Conversely, RNA-based approaches are attractive, given that (1) RNA reflects precisely what T cells express, (2) RNA-based assays are more sensitive given the larger number of RNA copies per cells relative to DNA, and (3) RNA is compatible with unique molecular identifiers (UMIs) to correct amplification and sequencing errors.[22] However, although no clear evidence has been provided to date, RNA is thought to bias clonotype quantification because of variations in TCR expression among cells.[8,23] To address this question, we sought to take advantage of scTCR-seq technology, which provides both TCR messenger RNA (mRNA) expression and clonotype frequencies in the same assay. We thus performed scTCR-seq on tumor-infiltrating lymphocytes (TILs) from three melanoma patients to evaluate how TCRs based on RNA expression and clonotype frequencies correlate, a clonotype being defined as a T cell expressing the same TCR at the protein level (Figure 1A). We first evaluated the variation of TCR expression among all cells and calculated the number of TCR mRNA molecules per cell for each patient. T cells expressed on average 10 TCRα and 21 TCRβ mRNA molecules (Figure 1B), in agreement with previous predictions.[24] Furthermore, while a large inter-cell variation was observed, TCR expression varied within 2-fold around the median in 70% of the cells (Figure 1B). We then wondered whether this difference was clonotype dependent. The analysis of each cell composing the 20 most frequent clonotypes showed that intra-clonotype TCR expression heterogeneity was similar to the inter-cell heterogeneity, indicating that TCR expression is not clonotype dependent (Figures 1C and S1A–S1D). We then evaluated how T cell states impacted TCR expression. We first ranked clonotypes according to their frequencies, the most frequent representing the clonally expanded, and looked at both distribution and average TCR expression. Data demonstrated that TCR expression was not related to the clonotype frequency (Figures S2A–S2D). We then used single-cell RNA sequencing (scRNA-seq) data to quantify TCR expression in naive, activated, memory, effector memory, and exhausted T cells. Only minor changes of the average TCR expression were observed between the different T cell states, with a trend for the activated T cells to express fewer TCRs (Figures S2E and S2F). The lower amount of mRNA molecules detected with scRNA-seq as compared with scTCR-seq may reflect the better ability of scTCR-seq to capture TCRs owing to its specific amplification. Altogether, our data demonstrate that, even though we observed substantial variation of TCR expression between cells, this variation was not related to the TCR sequence or to the T cell states.

The above data suggested that inter-clonotype variation should remain minimally impacted. Indeed, if the TCR expression calculated to quantify clonotype frequencies is the result of multiple cells with a large range of TCR expression, the average expression will tend to the median and the inter-clonotype variation will be minimal. In agreement with this, most clonotypes (~80%) with an average TCR expression per cell exceeding a 2-fold range from the median were singletons, hence precluding

Figure 1. Variability of TCR expression levels

(A) Schematic description of single-cell TCR-seq analysis performed on TILs from three melanoma patients and the analysis of the number of TCR RNA molecules per cell.

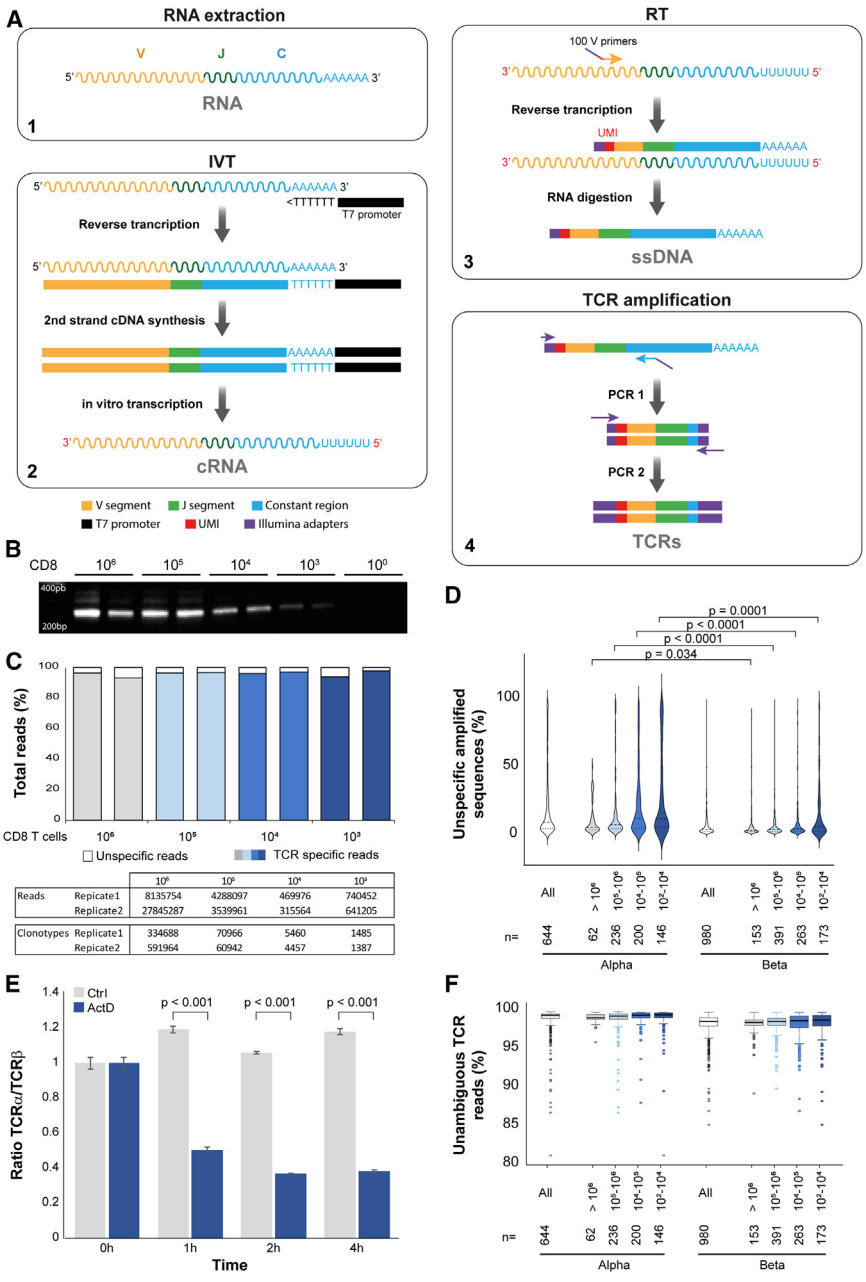
(B) Violin plots showing the distribution of the number of RNA molecules per cell normalized to the median for TCR α and β chains in three patients. The numbers in violin plots indicate the average mRNA molecules per cell for each patient. The number of single cells analyzed is indicated at the bottom of the graph.

(C) Heatmaps showing TCR mRNA expression levels of the 20 most frequent clonotypes (y axis). Each square represents one cell from each clonotype. Color scale highlights the number of mRNA molecules per cell. For readability, only 20 cells of the clonotype are presented. The TCR expression distribution of all the cells of the clonotypes are shown in Figure S1B.

(D) The average TCR expression levels (number of mRNA molecules per cell) were determined for each clonotype and plotted according to the number of cells per clonotype. Each dot shows an individual clonotype. Red lines represent median expression levels; dotted lines and gray zones represent the 2-fold variations from medians. The graph depicted here shows data of patient 1 and is representative of the three patients.

(E) All clonotypes with an average expression level of TCR mRNA >2-fold relative to medians (present in the gray zone of D) were deconvoluted according to the number of cells per clonotype.

(F and G) Correlations between TCR mRNA frequencies (y axis) and clonotype frequencies (number of cells/clonotype, x axis) based on single-cell data. p values and R^2 were calculated using Spearman correlation after logarithmic transformation of the data. The graphs depicted here show the data of patient 1. Data from patients 2 and 3 are shown in Figures S1E and S1F.



		10^6	10^5	10^4	10^3
Reads	Repicate1	8135754	4288097	469976	740452
	Repicate2	27845287	3539961	315564	641205
Clonotypes	Replicate1	334688	70966	5460	1485
	Repicate2	591964	60942	4457	1387

Figure 2. Validation of SEQTR

(A) Illustration of the different steps for TCR amplification using SEQTR. Sinusoidal lines represent RNA molecules and straight lines DNA or cDNA molecules.

(B) The amplified TCRs were run on an agarose gel to visualize the bands. The amplification was performed in duplicates with decreasing amounts of starting T cells.

(C) Frequencies of unspecific sequences (i.e., reads that do not align against TCR sequences) obtained after sequencing analyses of the repertoires amplified in (B). The table below summarizes the number of reads and the number of clonotypes identified in each sample.

(D) Violin plot showing the percentages of unspecific sequences identified in TCR α and β repertoires displayed according to the number of cells analyzed (x axis). Numbers at the bottom indicate the samples sequenced for repertoires analysis. p values were calculated by two-tailed t tests.

(E) RNA from PBMCs treated or not with actinomycin D was extracted at different time points, and TCRα/β RNA were quantified by real-time PCR. The graph shows TCRα/TCRβ RNA ratio. Each experiment was run in triplicate; the data shown are representative of the three independent experiments performed. p values were calculated by two-tailed unpaired t tests.

(F) The whisker plot shows the percentage of ambiguous V/J sequences (i.e., reads where several V or J segments are possible based on the sequencing and cannot be differentiated) identified in TCRα/β repertoires. Boxplots and box limits represent medians (line) and 25%–75% confidence limit, respectively. Whiskers are calculated as $Q_3 \pm 1.5\times$ interquartile range (IQR). Individual dots show outliers. Data were grouped according to the amount of starting cells. Numbers at the bottom indicate the number of repertoire samples analyzed.

any TCR expression average effect (Figures 1D and 1E). To finally legitimize the usage of RNA-based methods to accurately quantify clonotypes, we demonstrated that the inferred TCR frequencies based on RNA expression strongly correlated with the cognate clonotype frequencies (Figures 1F, 1G, S1E, and S1F).

SEQTR assay validation

The ultimate aim of TCR-seq assays is to yield quantitatively correct results, with an accurate relative ratio of TCRs within a given sample. To avoid bias, the two amplification steps used by SEQTR are (1) an *in vitro* transcription (IVT)[25] and (2) a PCR with one unique primer pair (Figure 2A). Of note, all amplicons have roughly the same size to avoid size-amplification biases. After extraction, mRNA is amplified by IVT and the resulting complementary RNA (cRNA) is reverse transcribed with a library of 100 V primers containing UMI and an Illumina adapter. Adding the collection of V primers here reduces the bias, as no amplification is done at this step. Purified single-strand DNA is then used to amplify TCRs with one primer in the Illumina adapter and one in the constant region. Finally, the different indexes required for sample multiplexing are added by a second round of PCR.

To assess specificity and sensitivity of SEQTR, TCRβ chains were amplified from 10^3 to 10^6 peripheral blood mononuclear cell (PBMC)-purified CD8 T cells. Amplicons were detected in all conditions, and unspecific sequences remained <6.5% even with only 10^3 T cells (Figures 2B and 2C). To further validate specificity, TCRα repertoires of 644 distinct sequenced samples and TCRβ repertoires of 980 distinct sequenced samples (TILs or

CellPress
OPEN ACCESS

A

	SEQTR			ImmunoSEQ		
	Replicate1	Replicate2	Replicate3	Replicate1	Replicate2	Replicate3
Total reads	9514300	8964593	10003171	NA	NA	NA
Aligned reads	9102193 (95.7%)	8602091 (95.9%)	9615107 (96.1%)	NA	NA	NA
Productive reads	8879501 (93.3%)	8397997 (93.6%)	9389792 (93.8%)	NA	NA	NA
Clonotypes	195308	164939	178164	155546	141276	139423

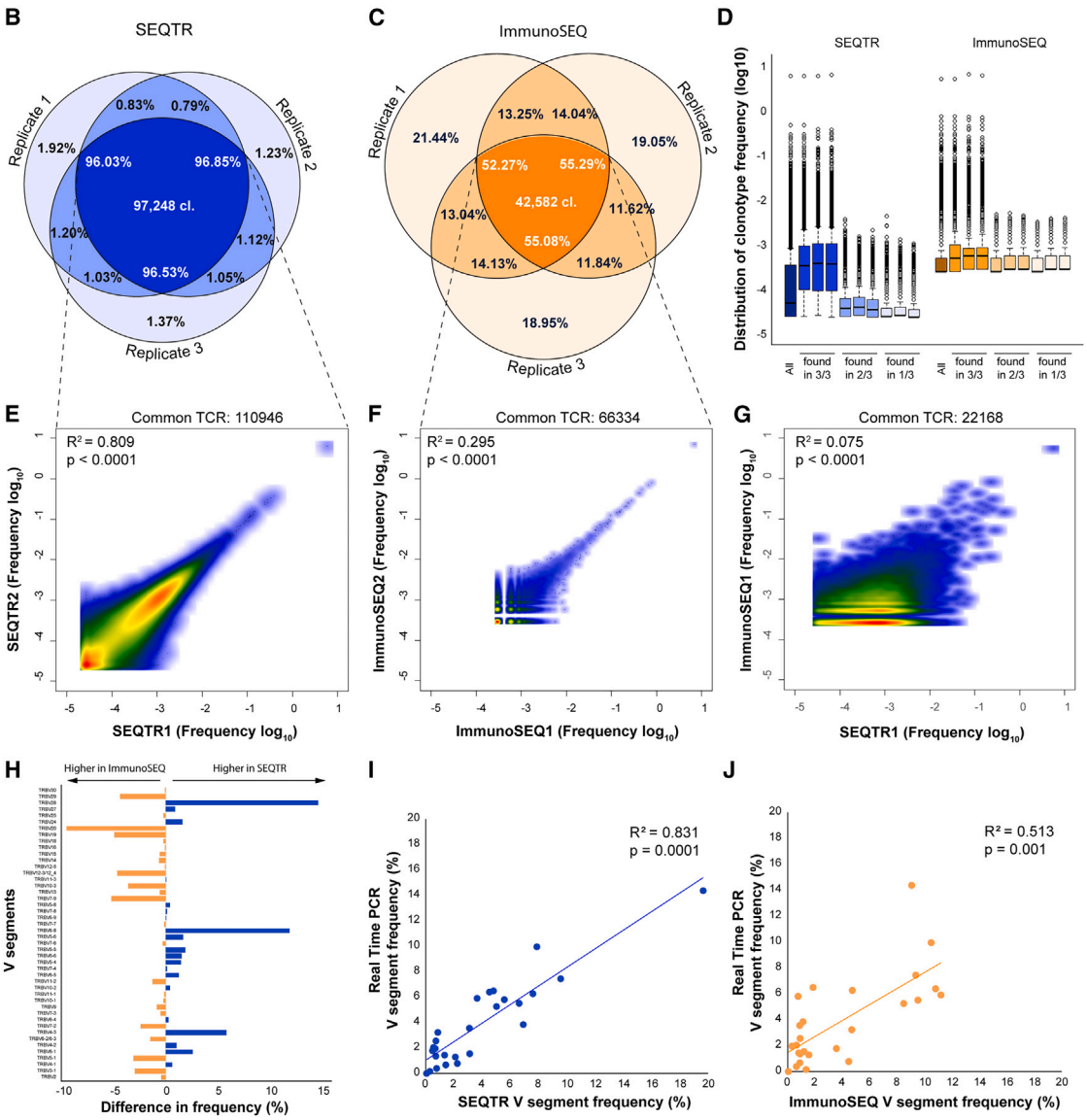


Figure 3. Reproducibility and sensitivity of SEQTR

RNA from 10^6 PBMCs (donor1) was extracted and used to perform three technical replicates with either SEQTR or ImmunoSEQ.

(A) The table summarizes the number of reads performed for each replicate, the number of aligned reads (i.e., those containing TCR sequences) and the productive reads (i.e., those where complete and unambiguous TCR was identified). The number of unique clonotypes identified is also indicated.

(B and C) Venn diagrams showing frequencies of overlapping TCR repertoires (frequency of common clonotypes) for SEQTR (B) and ImmunoSEQ (C). The number indicates the number of clonotypes identified in the three replicates.

(D) Distribution of clonotype frequencies calculated from the triplicates for complete repertoires (all) and for the fractions of clonotypes found in 3/3, 2/3, or 1/3 triplicates. Boxplots and box limits represent medians (line) and 25%–75% confidence limit, respectively. Whiskers are calculated as $Q_3 \pm 1.5\times$ IQR. Individual dots show outliers.

(E and F) Density plot showing frequencies of shared clonotypes from replicates 1 and 2 of SEQTR (D) and ImmunoSEQ (E).

(legend continued on next page)

PBMCs) were analyzed, and a median of unspecific sequences of 8.9% and 3.7% were identified for α and β, respectively (Figure 2D). Of note, unspecific sequences were more abundant when cell numbers were low (i.e., $<10^3$ cells) or in isolated samples, most probably reflecting a lower amount or quality of RNA. Given the significantly higher fraction of non-specific sequences in TCRα relative to TCRβ repertoires (Figure 2D), we postulated that only a common mechanism could explain this consistent effect. We thus evaluated RNA stability of α and β chains using actinomycin-D-treated PBMCs. Quantification of TCRα/β RNA by real-time PCR confirmed a faster RNA degradation for TCRα relative to TCRβ chains, consistent with the higher unspecific amplification and the previously reported issues in amplifying TCRα chains[8] (Figure 2E).

Another limitation of repertoire analyses is the unambiguous identification of TCR sequences. Ambiguity mostly occurs when discrimination between distinct V segments is impossible. As some V sequences only differ by 1 bp,[26] amplification or sequencing errors can affect their identification. Analysis of repertoires with SEQTR showed that >95.8% of sequences were unambiguous, irrespective of cell numbers, sample origin, and the chain sequenced (Figure 2F).

SEQTR benchmarking

We next benchmarked SEQTR and compared it with the widely used ImmunoSEQ assay.[10] To assess reproducibility (assay precision), RNA obtained from 10^6 PBMCs (donor 1) was sequenced in triplicate with SEQTR and with ImmunoSEQ. Qualitatively, overlaps between TCRβ repertoire triplicates were 96% for SEQTR and 54% for ImmunoSEQ ($p < 0.0001$), yet more clonotypes were systematically detected with SEQTR (n = 97,248 vs. 42,582; Figures 3A–3C). Frequencies of the non-overlapping fractions (i.e., clonotypes not consistently captured in all triplicates) were significantly lower with SEQTR, further highlighting its higher sensitivity and reproducibility (Figures 3B–3D). Quantitatively, correlations between clonotype frequencies among duplicates were stronger for SEQTR (mean on n = 3: R^2 = 0.778 vs. 0.288 [Figures 3E and 3F] and 0.773 vs. 0.286 [Figures S3A and S3B] for an independent validation). Surprisingly, however, we found only a poor correlation (mean on n = 3: R^2 = 0.050 and 0.038) between SEQTR and ImmunoSEQ (Figures 3G and S3C), indicating that each assay estimated differently the frequency of shared clonotypes. To highlight potential bias introduced during amplification, T cell clones with known TCRs were spiked into irrelevant PBMCs at varying concentrations and then sequenced. Clonotype concentrations showed good linearity with both assays (Figure S3D). We reasoned that if such a systematic bias was introduced in all conditions during amplification, the ratios of diluted clonotypes would be maintained and that comparison of multiple TCRs was needed to observe any bias. As each method uses different collections of V primers during library construction, it could represent the main source of discrepancies between them. We thus compared the frequencies of the different V segments provided by repertoire analyses and found a clear difference (Figure 3H). Similar results were obtained when using a custom multiplex TCR amplification (Figures S3G–S3I). To evaluate which method was the most truthful, we quantified V segments in the bulk RNA sample by real-time PCR, whereby amplification biases related to primer efficiencies can be corrected, and compared their frequencies with those obtained by sequencing. Analysis of 35 out of 46 functional V segments revealed that SEQTR was quantitatively more reliable (R^2 = 0.831 and 0.741) than ImmunoSEQ (R^2 = 0.513 and 0.044; Figures 3I, 3J, S3E, and S3F). Finally, we calculated different TCR metrics on the replicate repertoires. The consistency of the metrics further highlighted the reproducibility of the two methods but also demonstrated significant differences between them (Figures 5B and 5C).

While most of the TCR repertoire studies use genomic DNA combined with a multiplex PCR assay, the 5′-RACE approach is the reference for RNA-based assays. Therefore, we also benchmarked SEQTR against SMARTer, a commercially available 5′-RACE assay. The potential limit of the 5′-RACE being the sensitivity, we tested the two methods on a highly diverse repertoire. RNA from 3×10^6 PBMCs was extracted and sequenced in triplicate with SEQTR or SMARTer. The first striking difference was the number of unproductive reads (i.e., those where no TCR were identified). Ninety-four percent of the reads was exploited with SEQTR, as opposed to only 61.4% with SMARTer (Figure 4A). The overlap between the replicate repertoires suggested that the reproducibility of SMARTer (27.9% overlap) was similar to that of SEQTR (33.7%) (Figures 4B and 4C). However, the number of clonotypes systematically detected was almost twice as low with SMARTer relative to SEQTR, in agreement with a lower sensitivity of the 5′-RACE approach. This reduction of sensitivity was also observed in the total number of clonotypes detected (mean on n = 3: 198,324 with SEQTR vs. 137,528 with SMARTer) (Figure 4A) and in the frequency distribution of the clonotypes identified (Figure 4D). The difference between SEQTR and SMARTer indeed demonstrated the reduction of low-frequency TCRs detected with SMARTer (Figure 4D). Quantitatively, correlations between clonotype frequencies among duplicates were only slightly higher for SEQTR (mean on n = 3: R^2 = 0.195 vs. 0.109 [Figures 4E and 4F] and R^2 = 0.255 vs. 0.129 [Figures S3J and S3K] for an independent validation). Similarly to the comparison with ImmunoSEQ, clonotype quantification between the two assays was different (Figures 4G and S3L) with reproducible variations in the V segment usage (Figure 4H), though less pronounced than between SEQTR and ImmunoSEQ. Quantification of the V segments in the starting RNA sample by real-time PCR showed a better correlation with SEQTR (R^2 = 0.714 and 0.560) than SMARTer (R^2 = 0.612 and 0.272) (Figures 4I, 4J, S3M, and S3N).

(G) Density plot showing frequencies of clonotypes shared between SEQTR and ImmunoSEQ. p value and R^2 were calculated using Spearman correlation after logarithmic transformation of the data. Only one representative comparison is shown in the figure.
(H) Frequencies of the different V segments were calculated for each method. The graph shows the difference in V frequencies between SEQTR and ImmunoSEQ.
(I and J) Cross-comparison between TCRβ-chain variable (TRBV) frequencies quantified by real-time PCR on the original RNA or by SEQTR (I) or ImmunoSEQ (J). V segments that could not be differentiated by real-time PCR were amplified with a common primer and corresponding sequencing frequencies pooled together. p values and R^2 were calculated using Spearman correlation.

A

	SEQTR			5'RACE (SMARTer)		
	Replicate1	Replicate2	Replicate3	Replicate1	Replicate2	Replicate3
Total reads	9737454	9290795	9088617	6493536	7139353	8910389
Aligned reads	9360967 (96.1%)	9010308 (97.0%)	8836985 (97.2%)	4843508 (74.59%)	4376562 (61.3%)	5607495 (62.9%)
Productive reads	9074784 (93.2%)	8752221 (94.2%)	8601456 (94.6%)	4545042 (70.0%)	4018626 (56.3%)	5175007 (58.1%)
Clonotypes	168290	211088	215595	142733	135057	134796

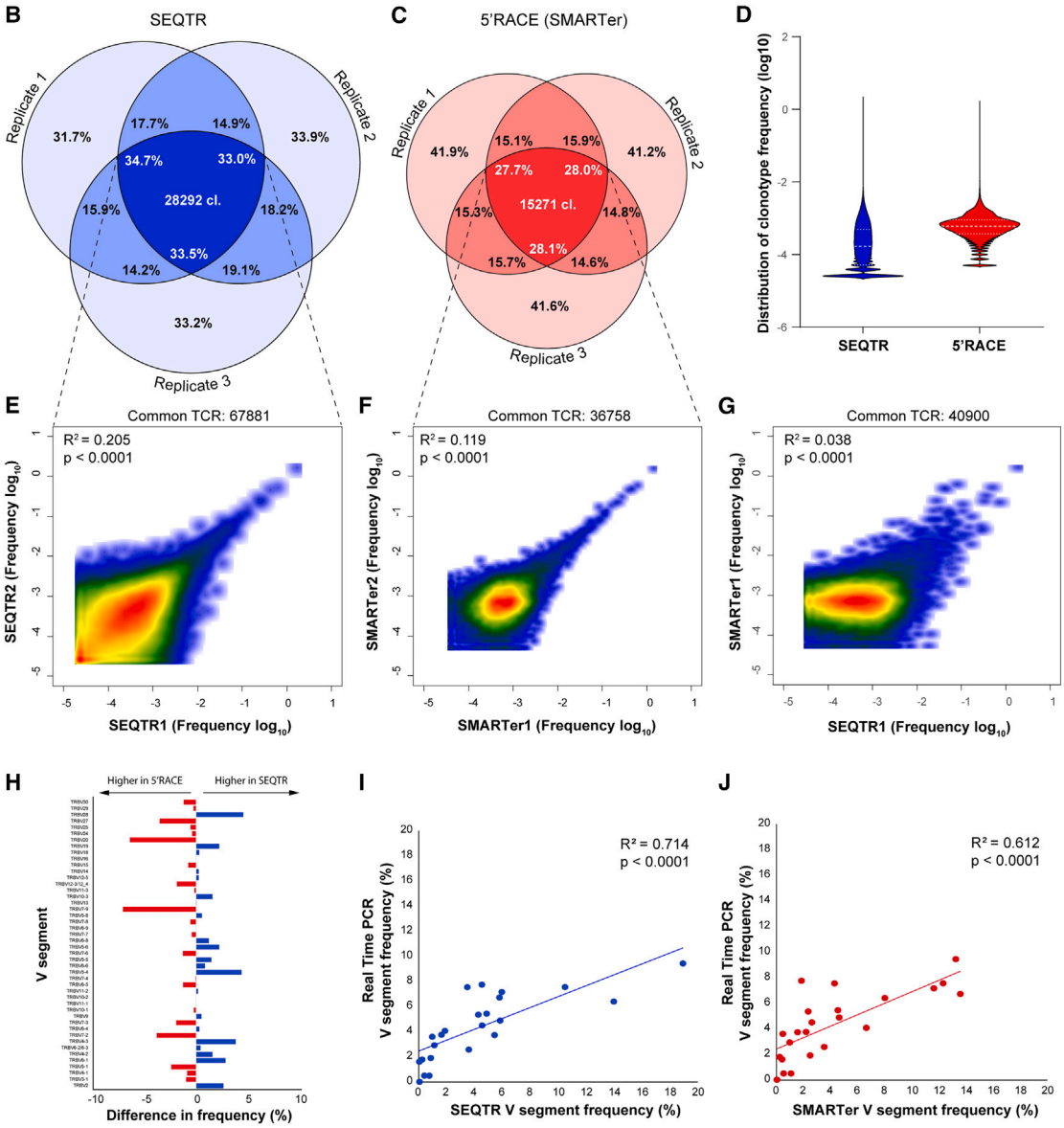


Figure 4. Reproducibility and sensitivity of SEQTR

RNA from 3 × 10^6 PBMCs (donor3) was extracted and used to perform three technical replicates with either SEQTR or the 5′-RACE assay SMARTer.

(A) The table summarizes the number of reads performed for each replicate, the number of aligned reads (i.e., those contained TCR sequences), and the productive reads (i.e., those where a complete and unambiguous TCR was identified). Number of unique clonotypes identified is also indicated.

(B and C) Venn diagrams showing frequencies of overlapping TCR repertoires (frequency of common clonotypes) for SEQTR (B) and SMATer (C). The number indicates the number of clonotypes identified in the three replicates.

(D) Violin plot showing the frequency distribution of all clonotypes identified. Only one representative replicate is represented.

(E and F) Density plots showing frequencies of shared clonotypes from replicates 1 and 2 of SEQTR (E) and SMARTer (F).

(G) Density plots showing frequencies of clonotypes shared between SEQTR and SMARTer. p value and R^2 were calculated using Spearman correlation after logarithmic transformation of the data. Only one representative comparison is shown in the figure.

(legend continued on next page)

To extend our observations, we next designed primers to amplify murine TCR α/β chains. SEQTR provided reproducible and accurate estimation of murine TCR repertoires, even though TCRα V segment identification was more challenging because of the high homology between replicated V segments (Figures S4 and S5).

Altogether, our data demonstrated that SEQTR is a quantitative, sensitive, and reproducible method to analyze human and mouse TCR repertoires. As expected, the multiplex and 5′-RACE approaches are both limited by an increasing quantification bias for the first one and a reduced sensitivity for the second. SEQTR combines sensitivity and accuracy and thus provides the most reliable representation of repertoires.

Impact of TCR analysis methods on metrics measurements

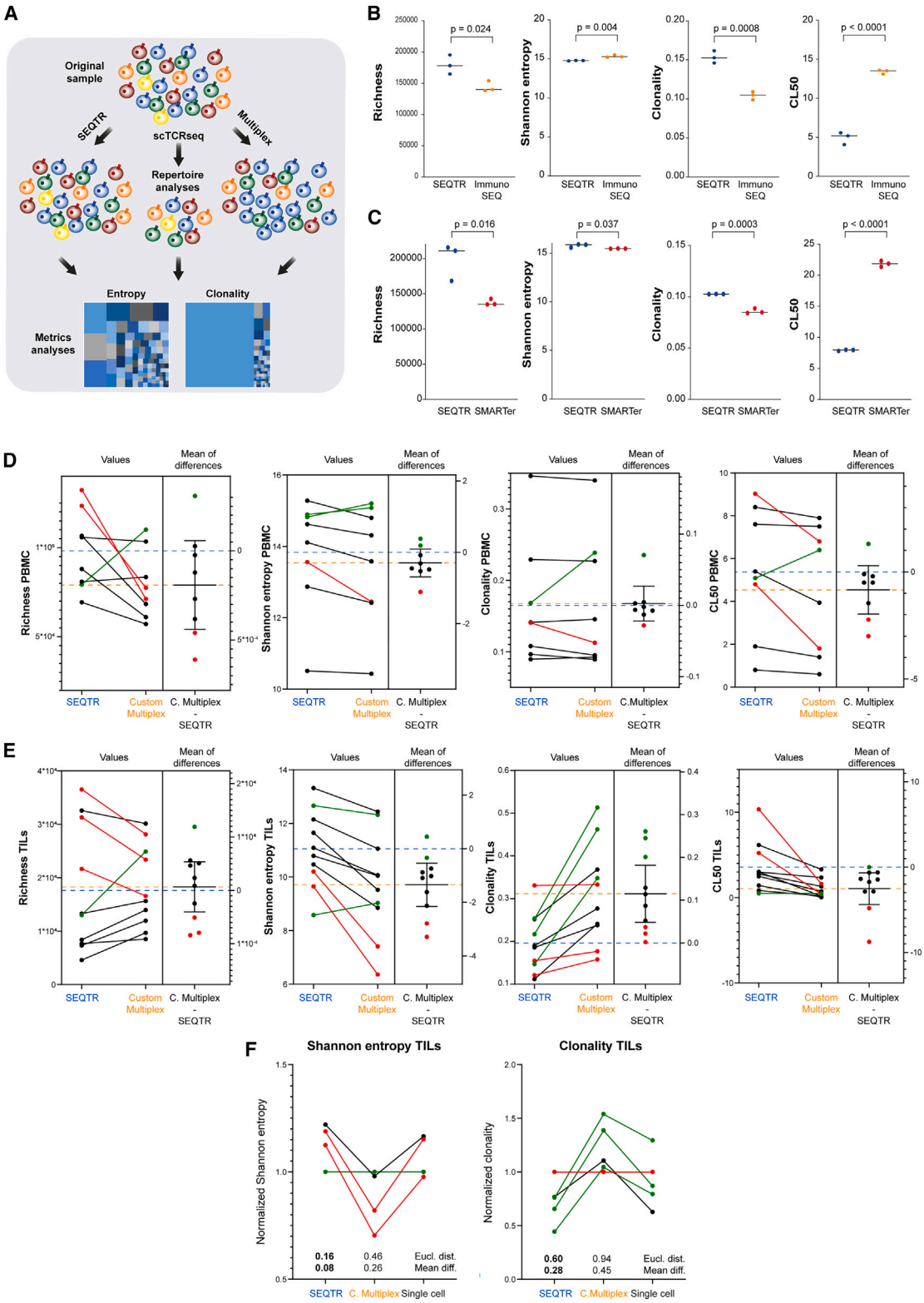
TCR repertoire metrics are frequently used to profile and monitor immune responses in inflammatory diseases, transplantation, vaccinology, and cancer immunotherapy.[1–4,27] Shannon entropy and clonality, which are metrics applied to determine repertoire richness and evenness, were recently established as biomarkers for patient stratification in cancer immunotherapies.[28–30] However, biases in clonotype quantification (evenness) or difference in methods sensitivity (richness) may lead to inaccurate representations of repertoires and thus to misleading metrics measurements, ultimately reducing the application of these metrics to clinical settings (Figure 5A). To evaluate how the accuracy of each method may affect metrics, we first calculated them on the replicates of PBMC samples used to benchmark SEQTR. For all metrics analyzed, significant differences appeared between SEQTR and ImmunoSEQ or SMARTer, confirming the impact of methods for metrics investigations (Figures 5B and 5C). To better understand this bias, eight PBMCs from healthy donors (Figure 5D) and ten TILs from melanoma patients (Figure 5E) were analyzed with both SEQTR and our custom multiplex assay. Using Bland-Altman analysis,[31] we evaluated the extent of discrepancy between the two methods. To this end, differences between SEQTR or multiplex assay measurements were calculated for each sample, and the limit of agreement was defined by the 95% confidence interval (CI) of the differences. The first striking observation was the inconsistent variation of measurements (Figures 5D and 5E). Indeed, for both PBMCs and TILs, the 95% CI variability was exceeded by several samples, demonstrating that amplification biases differently affect individual sample metrics. The quantitative and sensitive nature of SEQTR suggests that metrics calculated on a less biased repertoire would therefore be more accurate. To confirm this hypothesis, we used scTCR-seq data as a third orthogonal approach to profile samples. While single-cell analyses are not devoid of any bias and are not expected to strictly mimic bulk methods because of the lower number of sequenced cells, analyzing the metrics ratio may shed light on samples with large discrepancy between SEQTR and multiplex PCR. We selected four samples for Shannon entropy and five for clonality where significant differences were observed. Data showed that, after normalization, Shannon entropy and clonality obtained with scTCR-seq were more consistent and further supported SEQTR (Figure 5F). Together, our data highlight the importance of unbiased TCR repertoire analyses to fully exploit the value of TCR metrics in clinical settings. Indeed, while consistent biases between methods would not have impaired the relative sample ranking, inconsistent biases, as observed here, impair granular analyses and any downstream discovery or application.

We pushed the comparison between single-cell data and the bulk methods to further illustrate the lower bias induced by SEQTR. Owing to the low number of cells analyzed with scTCR-seq, we first concentrated on the most frequent TCRs and evaluated how many of the top 20 single-cell TCRs (scTCRs) were shared with the top 20 TCRs identified with SEQTR or the multiplex assay. Analyses of the 10 TIL samples showed a higher number of the top 20 scTCRs shared with the top 20 TCRs sequenced with SEQTR than the top 20 TCRs sequenced with the multiplex assay (Figure S6A). Indeed, 55 of the 200 top 20 scTCRs were found with SEQTR TCRs compared with 23 with the multiplex amplification (Figure S6B). When considering each patient individually, the number of TCRs shared among the top 20 TCRs was also significantly higher with SEQTR (Figure S6C). Despite the huge difference of cells analyzed between single-cell and bulk assays, we attempted to perform a global comparison of the repertoires. Interestingly, the correlation and the homology between scTCR repertoires and SEQTR were significantly higher than comparison between scTCR repertoires and the multiplex (Figures S6D–S6F). Altogether, these data confirm the higher accuracy of SEQTR over multiplex assays.

TCR cloning from bulk T cell material

While scTCR-seq represents an unprecedented tool to identify TCRs, cloning TCRs remains a mandatory step to validate their recognition of candidate targets or to engineer T cells for adoptive cell transfer therapies. Different methods have been developed to clone single TCR or TCR libraries.[32–34] However, those methods either fail to include SNP for individual sample or require a T cell cloning step to amplify the TCR of interest. We developed here a strategy allowing TCR amplification directly from the leftover of scTCR-seq libraries or from the remaining RNA from SEQTR, avoiding any cloning step and including all TCR SNPs (Figure 6A). To the best of our knowledge, full-length TCR amplification from bulk T cell material has never been reported, as V segments and constant regions are shared by many TCRs, thus making specific full-length TCR amplification impossible. To circumvent this issue, TCRs were amplified in two fragments: one containing the V and the CDR3 and one containing the CDR3, the J, and the constant region, while specificity was ensured by primers in CDR3. The two overlapping products were then combined by fusion PCR to amplify the variable regions of TCRs of interest (Figure 6A). Next, to allow fast TCR

(H) Frequencies of the different V segments were calculated for each method. The graph shows the difference in V frequencies between SEQTR and SMARTer.
(I and J) Cross-comparison between TRBV frequencies quantified by real-time PCR on the original RNA or by SEQTR (I) or SMARTer (J). V segments that could not be differentiated by real-time PCR were amplified with a common primer and corresponding sequencing frequencies pooled together. p values and R^2 were calculated using Spearman correlation.



(legend on next page)

cloning into vectors, variable regions can be mixed with three vectors containing the α constant, the β constant of choice, and the backbone, respectively, and combined in a single cloning step using the golden gate assembly (Figure 6A). Alternatively, electroporation with *in vitro*-transcribed TCR mRNA has been shown to be an efficient and rapid method to express functional TCRs on T cells.[35] To this end, a second fusion PCR is performed to combine variable regions with the T7 promoter and either the human constant region or alternative (e.g., murine) constant regions to reduce potential mispairing.[36] Therefore, the final PCR products can be immediately used for IVT, reducing the time required to clone and screen TCRs (Figure 6A).

To validate our strategy, using fluorescence-activated cell sorting (FACS) we purified HLA-A*0201 $CMV_{NLVPMVATV}$-specific $CD8^+$ T cells from PBMCs using peptide major histocompatibility complex multimers and sequenced their TCRs with SEQTR (Figures 6B and 6C). Unique dominant α and β variable regions were amplified from bulk RNA and combined with the mouse constant region. Human leukocyte antigen (HLA) class I matched PBMCs were retrovirally transduced with TCR and then stained with the cognate multimer. Following transduction, 77% of total $CD8^+$ T cells and 89% of transfected cells (carrying the mouse constant region) stained positively with the cognate multimer (Figure 6D). Similarly, TCRs were successfully amplified from the leftover of scTCR-seq libraries, demonstrating that the aforementioned fusion PCR approach can be used to amplify antigen-specific TCRs from bulk material directly downstream of SEQTR or scTCR-seq. Out of the 300 TCRs we attempted to clone, 285 (95%) were successfully amplified with our strategy.

Fast and reliable identification tumor-reactive TCRs

Here we combined the two aforementioned strategies to identify and validate clinically relevant tumor-reactive TCRs in a time- and cost-effective way. To this end, we exposed *in vitro*-expanded TILs to autologous tumor cell lines and FACS-sorted tumor-reactive ($CD137^+$) and non-reactive ($CD137^-$) $CD8^+$ TILs. We used SEQTR to sequence the tumor-reactive $CD137^+$ population and compared it with unstimulated TILs. Taking advantage of the quantitative nature of SEQTR, we paired, according to their relative frequencies, the enriched α and β chains (i.e., candidate tumor-specific TCRs) from $CD137^+$ repertoires (Figure 7A).

Using TILs from a representative melanoma patient, we found one dominant α and two dominant β chains in the repertoire (Figures 7B–7D). The clonally expanded variable regions were amplified according to the above method from bulk RNA and fused to the mouse constant region. TCRs were retrovirally transduced in autologous PBMCs and tested for autologous tumor cell recognition. The data showed that only the α1/β1 TCR was able to recognize tumor cells in an HLA-class-I-restricted fashion (Figure 7E). We validated the functional α/β pair resulting from the above process as naturally occurring TCRs in matched scTCR-seq datasets (Figure 7E). Furthermore, we confirmed the ability of T cells transduced with this TCR to control autologous tumor growth *in vivo* (Figure 7F). The low cost of bulk relative to single-cell sequencing makes SEQTR an attractive alternative to identify clinically relevant TCRs. To further validate the robustness of this workflow to isolate tumor-specific TCRs, two additional melanoma patients were analyzed. $CD137^+$ $CD8^+$ T cells were sequenced with SEQTR, and resulting repertoires were compared with that of unstimulated TILs. Repertoire analyses showed oligoclonal α and β chains among $CD137^+$ $CD8^+$ TILs from both patients (Figures S7A–S7D). We transduced different combinations of α and β chains in autologous PBMCs and tested the α/β pairs for autologous tumor cell recognition. In both patients, we found that α and β chains from both repertoires (i.e., α1/β2 or α2/β1 and α1/β1) were functionally paired and that the cognate α1/β1 TCRs were recognizing tumor cells in an HLA-class-I-restricted fashion (Figure S7E). We also validated these pairs in a matched scTCR-seq dataset (Figure S7E). Again, in patient 2 for whom we succeeded in establishing an autologous tumor line, we confirmed the ability of autologous T cells transduced with the functional private TCRs to control autologous tumor growth *in vivo* (Figure S7F).

Taken together, these data demonstrate that bona fide tumor-specific TCRs can be identified and cloned in a sensitive as well as time- and cost-effective way. If SEQTR requires bias repertoire (i.e., $CD137^+$ sorted cells) to successfully pair α/β chains based on their frequency, the cloning strategy described here can also be positioned downstream of scTCR-seq for samples with more diverse repertoire.

DISCUSSION

In the fast-expanding field of TCR repertoire analyses, data accuracy largely relies on the applied method. The major challenge is to avoid bias during TCR amplification to robustly capture repertoire diversity and quantify the different clonotypes. The widely used multiplex PCR is known to introduce biases in clonotype quantification related to differences in primer efficiency of

Figure 5. Impact of methods on TCR metrics

(A) Description of the impact of methods on TCR metrics measurement.

(B and C) TCR metrics were calculated on the replicate used to benchmark SEQTR. Analysis of the richness, Shannon entropy, clonality, and CL_{50} calculated from technical replicate repertoires (n = 3 replicates) obtained with SEQTR or ImmunoSEQ (B) or with SEQTR and SMARTer (C). Bars show medians. p values were calculated by two-tailed unpaired t tests.

(D and E) Eight PBMCs (D) and ten TILs (E) from melanoma patients were used for TCR-seq with SEQTR or our custom multiplex PCR amplification. TCR metrics were calculated, and the results obtained with the two methods are presented on estimation plots. The left part shows metrics data, and the right part shows the mean of differences between paired samples. Mean and the upper and lower 95% confidence intervals are shown on right panels. The dotted lines show means of SEQTR (blue) and of our custom multiplex assay (orange). Samples above and below the 95% confidence interval are highlighted in green and red, respectively. p values were calculated by two-tailed paired t tests.

(F) Metrics calculated with SEQTR, multiplex, or single-cell approach were normalized to the SEQTR sample with lowest Shannon entropy or highest clonality. Colors illustrate sample over or below the 95% confidence interval of (D). Euclidian distance and mean of differences were calculated between single-cell and SEQTR or single-cell and multiplex.

CellPress
OPEN ACCESS

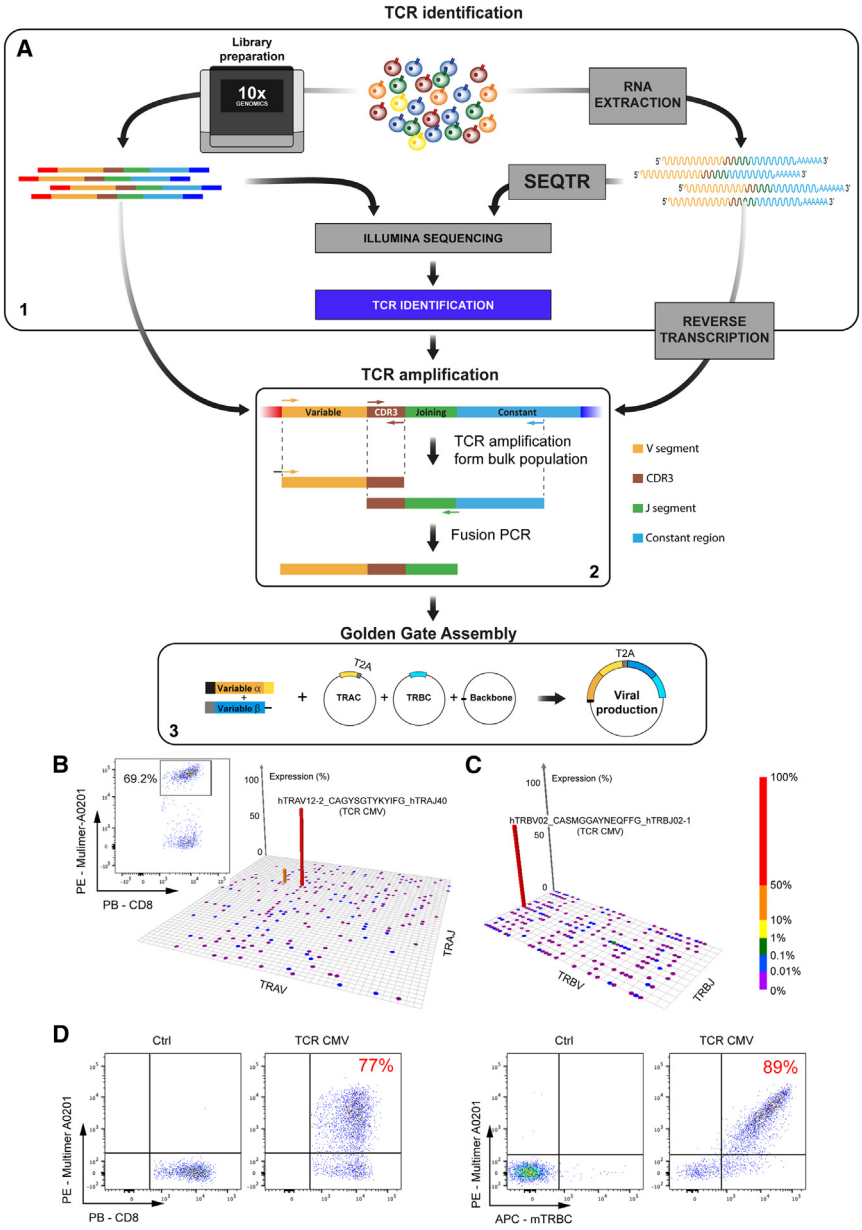


Figure 6. TCR cloning

(A) Illustration of the different steps to clone TCR from scTCR-seq library or bulk RNA.

(B and C) V and J recombinations for both TCRα (B) and TCRβ (C) chains are shown for cytomegalovirus (CMV)-tetramer-positive T cells (FACS plot in B). V/J segments are represented according to their chromosomal location on the x and y axis, respectively. The frequency of each recombination is shown on the z axis. The color code highlights recombination frequency.

(D) TCR-transduced T cells were stained with CD8 and tetramer loaded with CMV peptide (left panels) or CD8 and the mouse β constant region (right panels).

the multiple pairs.[14] While the 5′-RACE assay solved the multiplex issue by adding an anchor sequence in the 5′ end, efficiency of the template switch is quite poor with only 20%–60% of the cDNA correctly tagged,[16] resulting in lower sensitivity. If scTCR-seq has emerged as an efficient method to identify TCRs and link them to the physiological status of the cells, its cost prevents large-scale development of the method, and the lower number of cells analyzed renders scTCR-seq less appropriate when deeper analyses of repertoires are required. Here, we presented an approach which circumvents the aforementioned limitations and thus improves the quality of TCR repertoire analyses.

The pros and cons related to DNA or RNA as the source of material remains actively debated. The main drawback attributed to RNA is the lack of correspondence between TCR expression and clonotype frequency that would bias the analysis.[23] We confirmed substantial variations in TCR expression levels. However, this heterogeneity of TCR expression only marginally affects clonotype quantification, and RNA-based methods are suitable to evaluate TCR repertoires. The main reason is that intra-clonotype variation leads to an average TCR expression level that is consistent across different clonotypes. Nevertheless, 30% of low-frequency TCRs, and in particular singletons, could be affected by RNA expression, and this could potentially impact analyses of a highly diverse repertoire such as PBMCs. However, we showed that all methods have reached their limits to accurately quantify low-frequency TCRs. While we observed a larger variation of the reproducibility with SEQTR, such low-frequency TCRs are either not identified with ImmunoSEQ or poorly detected with 5′-RACE. Therefore, variation of TCR expression has more limited impact than the technical limits of detection of currently available methods. Still, RNA may not be appropriate for samples where RNA integrity is damaged. For example, in formalin-fixed paraffin-embedded samples, RNA is degraded by the tissue fixation and is thus incompatible with TCR amplification. In such cases, DNA remains the only alternative. In turn, whenever RNA degradation can be prevented, RNA-based methods would benefit from a higher sensitivity and a granular measurement of what T cells are expressing following allelic exclusion without compromising clonotype quantification.

While sensitivity is an important parameter enabled by RNA-based assays, the quantitative aspect of the method is even more fundamental. Indeed, while sensitivity allows analysis of samples of limited size and the detection of larger amounts of low-frequency TCRs, bias in quantification likely affects the

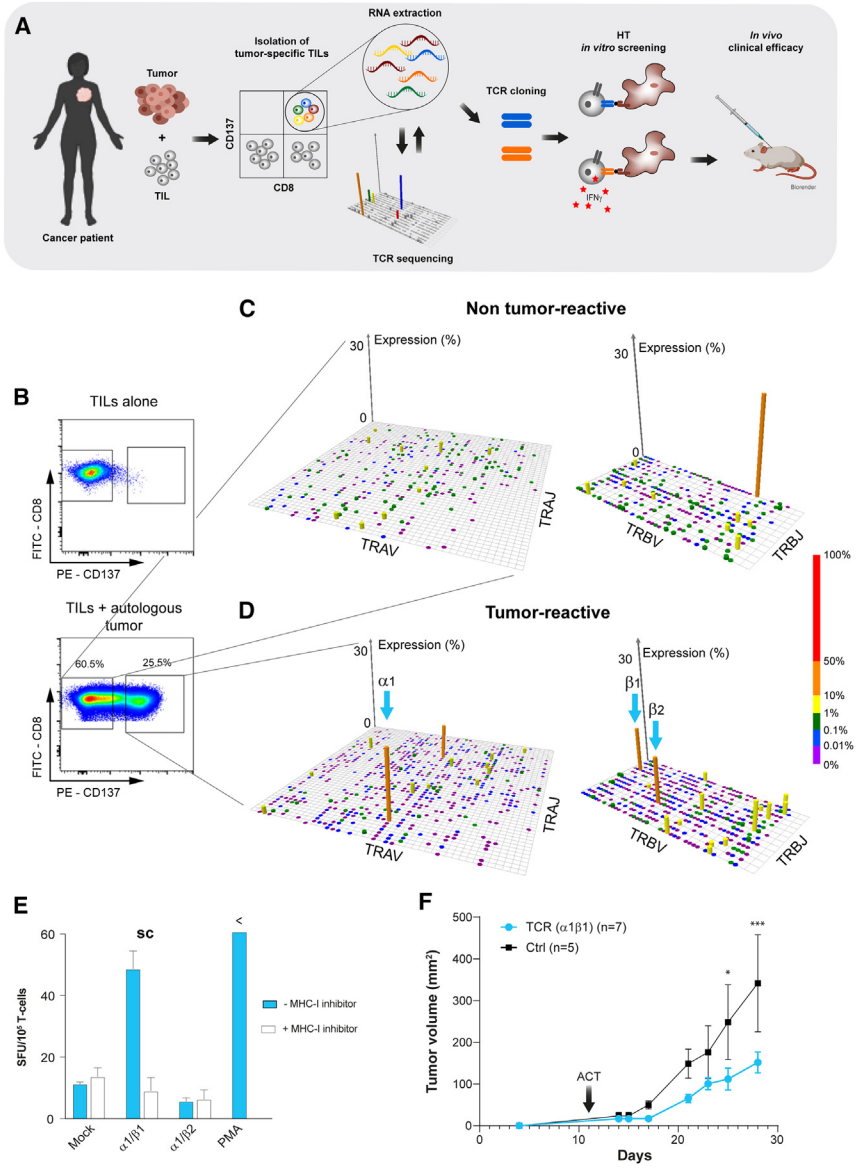


Figure 7. Identification of tumor-reactive TCRs from TILs

(A) Description of the TCR identification process. Tumor-reactive bulk TILs (CD137$^+$ after tumor recognition) were purified and sequenced. The most frequent TCRs were amplified from bulk RNA and tested *in vitro* for tumor recognition. Functional TCRs were tested for tumor control *in vivo*.

(B) Representative example of CD137$^+$ TILs alone or after 6 h of exposure to autologous tumor cells (patient 1).

(C and D) V/J recombinations for both TCRα and TCRβ chains are shown for non-tumor-reactive (CD137$^-$) (C) and tumor-reactive (CD137$^+$) (D) repertoires. V/J segments are represented according to their chromosomal location on the x and y axis, respectively. The frequency of each recombination is shown on the z axis. The color code highlights recombination frequency.

(E) Relative magnitude of tumor reactivity of transduced PBMCs with candidate TCRα/β chains from (C) and (D) measured by IFN-γ ELISpot assay in the absence (filled columns) or presence (open columns) of anti-HLA-class-I blocking antibody. "SC" indicates TCRs (cognate α- and β-chain pairs) found within matching single-cell TCR-seq data. Mean values and standard deviation (SD) are shown.

(F) Tumor control of adoptively transferred α1β1 TCR-transduced T cells against autologous patient 1-derived tumor xenografts. Mean values and SD are shown; p values were calculated by tow-tailed unpaired t tests.

entire repertoire including dominant clonotypes. It may thus completely misrepresent repertoire profiles and distort their interpretation. To avoid such limitations, we developed an approach based on two unbiased amplification steps: IVT was shown to assure linear amplification[37] and usage of a single primer pair that ensures the same amplification efficiency for all TCRs. Furthermore, adding the 5′ anchor sequence using reverse transcription has proven to be more efficient than the template switch, as shown by the higher performance of SEQTR in capturing low-frequency TCRs. However, it has to be noted that the sensitivity and reproducibility of the methods are also highly dependent on the starting material. Indeed, the RNA quality, the sample diversity, and the sequencing deepness will all impact assay reproducibility. High RNA degradation or sample diversity will affect assay reproducibility. Therefore, by its technology, SEQTR preserves repertoire integrity and provides data of greater accuracy and sensitivity than the multiplex PCR and 5′-RACE assays.

In the context of cancer immunotherapy, adoptive cell therapy and T cell engineering with tumor-specific TCRs have become attractive approaches.[6,7,38] Historically, TCR amplification biases represented major obstacles for the identification of functional α/β-chain pairs and thus also for the downstream cloning of tumor-specific TCRs.[39] The development of scTCR technologies[40,41] unleashed TCR α/β-chain pairing whereby emulsion-based technologies enabled high-throughput analyses.[42,43] If the challenge of TCR pairing has been solved, time- and cost-effective options to screen and clone TCRs are still needed. The solution provided here fulfills these requirements, allowing the amplification, cloning, and validation of a high number of TCRs in a couple of days. To achieve this, our pipeline relies on the ability to amplify TCRs directly from a bulk population which suppresses the delay required to either isolate the T cell clones or to synthesize DNA coding for TCRs. While ordering of the CD3-specific primers remains the limiting step of our strategy, the delays encountered to receive oligo were usually shorter than those for full-length TCR syntheses and could be reduced by benchtop oligo synthesizers. In addition, to reduce cloning time, direct amplification from sample material allows the

incorporation of all potential SNPs that may affect specificity and affinity for the cognate epitopes.[44] Combined with scTCR-seq, our method permits the identification and cloning of specific TCRs in a couple of days. However, the cost of single-cell analyses may remain an obstacle for high-throughput applications. Having demonstrated that our cloning strategy can also be positioned downstream of SEQTR, the combination of our two methods therefore represents an attractive and efficient opportunity to reduce the cost of the assay. Nevertheless, using bulk sequencing to paired α/β chains based on their frequencies is limited to samples with biased repertoires where dominant α or β chains can be identified.

With the development of immunotherapies, TCR analyses have generated an increasing interest and are now commonly used in various medical fields including cancer, infectious and autoimmune diseases, and transplantation.[45,46] Different tools have been developed to monitor and profile immune responses or to engineer T cells for adoptive transfer. For instance, TCR metrics were recently used as potential biomarkers to stratify patients and predict responses to immune checkpoint blockade therapy of cancer patients.[28–30,47] Similarly, the tracking of alloreactive clonotypes after organ transplantation may be used to adjust immunosuppressive therapies, diagnose or predict rejection, and monitor responses to anti-rejection therapies.[48,49] Successful applications of these TCR analyses will certainly rely on our ability to faithfully capture the complexity of repertoires to avoid misleading conclusions and unappropriated treatments and to provide time- and cost-effective tools to manipulate and engineer T cells. Thus, we believe that the methods presented here will be, by their reliability, robustness, and cost, valuable in releasing the full clinical potential of such therapies, not only for cancer immunotherapies as presented here but for any immune-related field.

Limitations of the study

SEQTR is a TCR repertoire analysis method based on RNA. Therefore, the success of the amplification is conditioned by RNA quality. In samples with partially or fully degraded RNA, amplification would be biased.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Patient samples
  - Mice model
- METHOD DETAILS
  - RNA extraction
  - TCR amplification: SEQTR (SEQuencing T cell Receptor)
  - Multiplex PCR
  - 5′ RACE amplification
  - TCR sequencing and analysis
  - TCR metrics
  - Single-cell transcriptome and TCR sequencing
  - Real-time PCR
  - Full length TCR cloning
  - T cell electroporation and tumor cell co-culture
  - Autologous tumor recognition by flow cytometry
  - T cell transduction and adoptive transfer in immunodeficient IL-2 NOG mice
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.crmeth.2023.100459.

ACKNOWLEDGMENTS

We are grateful to the patients for their dedicated collaboration and to healthy donors for their blood donations. We thank the staff of the CTE biobank for their assistance. We thank the Dr. Giacomo Oliveira for kindly sharing scRNA-seq data. This work was principally supported by grants from Cancera and Mats Paulssons, and by a gift from the Biltema Foundation (the Netherlands) that was administered by the ISREC Foundation (Switzerland).

AUTHOR CONTRIBUTIONS

R.G. and S.B. designed and performed experiments and analyzed data. J. Chiffelle, M.A., R.P., L.Q., A.M., P.R., J. Cesbron, A.A., P.B., P.G., and J.S. assisted with experimental work. P.R. assisted with experimental design. M.I., D.E.S., and L.E.K. provided additional support. R.G., G.C., and A.H. designed and supervised the study and wrote the manuscript.

DECLARATION OF INTERESTS

The University of Lausanne and Ludwig Institute for Cancer Research have filed for patent protection on the technology described herein. R.G. is named as inventor on this patent. G.C. has received grants from Celgene, Boehringer-Ingelheim, Roche, Bristol Myers Squibb, Iovance Therapeutics, and Kite Pharma. The institution with which G.C. is affiliated has received fees for G.C.'s participation on advisory boards or for presentation at a company-sponsored symposium from Genentech, Roche, Bristol Myers Squibb, AstraZeneca, NextCure, Geneos Tx, and Sanofi/Avensis. The Center Hospitalier Universitaire Vaudois (CHUV) and the Ludwig Institute for Cancer Research have filed for patent protection on the technology related to T cell expansion. S.B., A.H., and G.C. are named as inventors on this patent. G.C. has patents in the domain of antibodies and vaccines targeting the tumor vasculature as well as technologies related to T cell engineering for T cell therapy. G.C. holds patents around antibodies and receives royalties from the University of Pennsylvania regarding technology licensed to Novartis.

Received: October 17, 2022
Revised: February 27, 2023
Accepted: March 27, 2023
Published: April 24, 2023

REFERENCES

1. Babel, N., Stervbo, U., Reinke, P., and Volk, H.D. (2019). The identity card of T cells-clinical utility of T-cell receptor repertoire analysis in transplantation. Transplantation *103*, 1544–1555. https://doi.org/10.1097/TP.0000000000002776.
2. Li, N., Yuan, J., Tian, W., Meng, L., and Liu, Y. (2020). T-cell receptor repertoire analysis for the diagnosis and treatment of solid tumor: a

methodology and clinical applications. Cancer Commun. *40*, 473–483. https://doi.org/10.1002/cac2.12074.

3. Jacobsen, L.M., Posgai, A., Seay, H.R., Haller, M.J., and Brusko, T.M. (2017). T cell receptor profiling in type 1 diabetes. Curr. Diab. Rep. *17*, 118. https://doi.org/10.1007/s11892-017-0946-4.
4. Woodsworth, D.J., Castellarin, M., and Holt, R.A. (2013). Sequence analysis of T-cell repertoires in health and disease. Genome Med. *5*, 98. https://doi.org/10.1186/gm502.
5. Arnaud, M., Duchamp, M., Bobisse, S., Renaud, P., Coukos, G., and Harari, A. (2020). Biotechnologies to tackle the challenge of neoantigen identification. Curr. Opin. Biotechnol. *65*, 52–59. https://doi.org/10.1016/j.copbio.2019.12.014.
6. Chandran, S.S., Ma, J., Klatt, M.G., Dündar, F., Bandlamudi, C., Razavi, P., Wen, H.Y., Weigelt, B., Zumbo, P., Fu, S.N., et al. (2022). Immunogenicity and therapeutic targeting of a public neoantigen derived from mutated PIK3CA. Nat. Med. *28*, 946–957. https://doi.org/10.1038/s41591-022-01786-3.
7. Leidner, R., Sanjuan Silva, N., Huang, H., Sprott, D., Zheng, C., Shih, Y.P., Leung, A., Payne, R., Sutcliffe, K., Cramer, J., et al. (2022). Neoantigen T-cell receptor gene therapy in pancreatic cancer. N. Engl. J. Med. *386*, 2112–2119. https://doi.org/10.1056/NEJMoa2119662.
8. Barennes, P., Quiniou, V., Shugay, M., Egorov, E.S., Davydov, A.N., Chudakov, D.M., Uddin, I., Ismail, M., Oakes, T., Chain, B., et al. (2021). Benchmarking of T cell receptor repertoire profiling methods reveals large systematic biases. Nat. Biotechnol. *39*, 236–245. https://doi.org/10.1038/s41587-020-0656-3.
9. Krangel, M.S. (2009). Mechanics of T cell receptor gene rearrangement. Curr. Opin. Immunol. *21*, 133–139. https://doi.org/10.1016/j.coi.2009.03.009.
10. Robins, H.S., Campregher, P.V., Srivastava, S.K., Wacher, A., Turtle, C.J., Kahsai, O., Riddell, S.R., Warren, E.H., and Carlson, C.S. (2009). Comprehensive assessment of T-cell receptor beta-chain diversity in alphabeta T cells. Blood *114*, 4099–4107. https://doi.org/10.1182/blood-2009-04-217604.
11. Freeman, J.D., Warren, R.L., Webb, J.R., Nelson, B.H., and Holt, R.A. (2009). Profiling the T-cell receptor beta-chain repertoire by massively parallel sequencing. Genome Res. *19*, 1817–1824. https://doi.org/10.1101/gr.092924.109.
12. Douek, D.C., Betts, M.R., Brenchley, J.M., Hill, B.J., Ambrozak, D.R., Ngai, K.L., Karandikar, N.J., Casazza, J.P., and Koup, R.A. (2002). A novel approach to the analysis of specificity, clonality, and frequency of HIV-specific T cell responses reveals a potential mechanism for control of viral escape. J. Immunol. *168*, 3099–3104. https://doi.org/10.4049/jimmunol.168.6.3099.
13. Wang, C., Sanders, C.M., Yang, Q., Schroeder, H.W., Jr., Wang, E., Babrzadeh, F., Gharizadeh, B., Myers, R.M., Hudson, J.R., Jr., Davis, R.W., and Han, J. (2010). High throughput sequencing reveals a complex pattern of dynamic interrelationships among human T cell subsets. Proc. Natl. Acad. Sci. USA *107*, 1518–1523. https://doi.org/10.1073/pnas.0913939107.
14. Polz, M.F., and Cavanaugh, C.M. (1998). Bias in template-to-product ratios in multitemplate PCR. Appl. Environ. Microbiol. *64*, 3724–3730.
15. Carlson, C.S., Emerson, R.O., Sherwood, A.M., Desmarais, C., Chung, M.W., Parsons, J.M., Steen, M.S., LaMadrid-Herrmannsfeldt, M.A., Williamson, D.W., Livingston, R.J., et al. (2013). Using synthetic templates to design an unbiased multiplex PCR assay. Nat. Commun. *4*, 2680. https://doi.org/10.1038/ncomms3680.
16. Wulf, M.G., Maguire, S., Humbert, P., Dai, N., Bei, Y., Nichols, N.M., Corrêa, I.R., Jr., and Guan, S. (2019). Non-templated addition and template switching by Moloney murine leukemia virus (MMLV)-based reverse transcriptases co-occur and compete with each other. J. Biol. Chem. *294*, 18220–18231. https://doi.org/10.1074/jbc.RA119.010676.
17. Dahal-Koirala, S., Balaban, G., Neumann, R.S., Scheffer, L., Lundin, K.E.A., Greiff, V., Sollid, L.M., Qiao, S.W., and Sandve, G.K. (2022). TCRpower: quantifying the detection power of T-cell receptor sequencing with a novel computational pipeline calibrated by spike-in sequences. Brief. Bioinform. *23*, bbab566. https://doi.org/10.1093/bib/bbab566.
18. Li, S., Sun, J., Allesøe, R., Datta, K., Bao, Y., Oliveira, G., Forman, J., Jin, R., Olsen, L.R., Keskin, D.B., et al. (2019). RNase H-dependent PCR-enabled T-cell receptor sequencing for highly specific and efficient targeted sequencing of T-cell receptor mRNA for single-cell and repertoire analysis. Nat. Protoc. *14*, 2571–2594. https://doi.org/10.1038/s41596-019-0195-x.
19. Lin, Y.H., Hung, S.J., Chen, Y.L., Lin, C.H., Kung, T.F., Yeh, Y.C., Tseng, J.T., and Liu, T. (2020). Dissecting efficiency of a 5' rapid amplification of cDNA ends (5'-RACE) approach for profiling T-cell receptor beta repertoire. PLoS One *15*, e0236366. https://doi.org/10.1371/journal.pone.0236366.
20. Bassing, C.H., Swat, W., and Alt, F.W. (2002). The mechanism and regulation of chromosomal V(D)J recombination. Cell *109*, S45–S55. https://doi.org/10.1016/s0092-8674(02)00675-x.
21. Brady, B.L., Steinel, N.C., and Bassing, C.H. (2010). Antigen receptor allelic exclusion: an update and reappraisal. J. Immunol. *185*, 3801–3808. https://doi.org/10.4049/jimmunol.1001158.
22. Shugay, M., Britanova, O.V., Merzlyak, E.M., Turchaninova, M.A., Mamedov, I.Z., Tuganbaev, T.R., Bolotin, D.A., Staroverov, D.B., Putintseva, E.V., Plevova, K., et al. (2014). Towards error-free profiling of immune repertoires. Nat. Methods *11*, 653–655. https://doi.org/10.1038/nmeth.2960.
23. Nielsen, S.C.A., and Boyd, S.D. (2018). Human adaptive immune receptor repertoire analysis-Past, present, and future. Immunol. Rev. *284*, 9–23. https://doi.org/10.1111/imr.12667.
24. Ma, K.Y., He, C., Wendel, B.S., Williams, C.M., Xiao, J., Yang, H., and Jiang, N. (2018). Immune repertoire sequencing using molecular identifiers enables accurate clonality discovery and clone size quantification. Front. Immunol. *9*, 33. https://doi.org/10.3389/fimmu.2018.00033.
25. Moll, P.R., Duschl, J., and Richter, K. (2004). Optimized RNA amplification using T7-RNA-polymerase based in vitro transcription. Anal. Biochem. *334*, 164–174. https://doi.org/10.1016/j.ab.2004.07.013.
26. Lefranc, M.P. (2014). Immunoglobulin and T Cell receptor genes: IMGT((R)) and the birth and rise of immunoinformatics. Front. Immunol. *5*, 22. https://doi.org/10.3389/fimmu.2014.00022.
27. Chiffelle, J., Genolet, R., Perez, M.A., Coukos, G., Zoete, V., and Harari, A. (2020). T-cell repertoire analysis and metrics of diversity and clonality. Curr. Opin. Biotechnol. *65*, 284–295. https://doi.org/10.1016/j.copbio.2020.07.010.
28. Hogan, S.A., Courtier, A., Cheng, P.F., Jaberg-Bentele, N.F., Goldinger, S.M., Manuel, M., Perez, S., Plantier, N., Mouret, J.F., Nguyen-Kim, T.D.L., et al. (2019). Peripheral blood TCR repertoire profiling may facilitate patient stratification for immunotherapy against melanoma. Cancer Immunol. Res. *7*, 77–85. https://doi.org/10.1158/2326-6066.CIR-18-0136.
29. Hopkins, A.C., Yarchoan, M., Durham, J.N., Yusko, E.C., Rytlewski, J.A., Robins, H.S., Laheru, D.A., Le, D.T., Lutz, E.R., and Jaffee, E.M. (2018). T cell receptor repertoire features associated with survival in immunotherapy-treated pancreatic ductal adenocarcinoma. JCI Insight *3*, e122092. https://doi.org/10.1172/jci.insight.122092.
30. Kidman, J., Principe, N., Watson, M., Lassmann, T., Holt, R.A., Nowak, A.K., Lesterhuis, W.J., Lake, R.A., and Chee, J. (2020). Characteristics of TCR repertoire associated with successful immune checkpoint therapy responses. Front. Immunol. *11*, 587014. https://doi.org/10.3389/fimmu.2020.587014.
31. Bland, J.M., and Altman, D.G. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. Lancet *1*, 307–310.
32. Hu, Z., Anandappa, A.J., Sun, J., Kim, J., Leet, D.E., Bozym, D.J., Chen, C., Williams, L., Shukla, S.A., Zhang, W., et al. (2018). A cloning and expression system to probe T-cell receptor specificity and assess

functional avidity to neoantigens. Blood *132*, 1911–1921. https://doi.org/10.1182/blood-2018-04-843763.

33. Kato, T., Matsuda, T., Ikeda, Y., Park, J.H., Leisegang, M., Yoshimura, S., Hikichi, T., Harada, M., Zewde, M., Sato, S., et al. (2018). Effective screening of T cells recognizing neoantigens and construction of T-cell receptor-engineered T cells. Oncotarget *9*, 11009–11019. https://doi.org/10.18632/oncotarget.24232.
34. Wälchli, S., Løset, G.Å., Kumari, S., Johansen, J.N., Yang, W., Sandlie, I., and Olweus, J. (2011). A practical approach to T-cell receptor cloning and expression. PLoS One *6*, e27930. https://doi.org/10.1371/journal.pone.0027930.
35. Zhao, Y., Zheng, Z., Cohen, C.J., Gattinoni, L., Palmer, D.C., Restifo, N.P., Rosenberg, S.A., and Morgan, R.A. (2006). High-efficiency transfection of primary human and mouse T lymphocytes using RNA electroporation. Mol. Ther. *13*, 151–159. https://doi.org/10.1016/j.ymthe.2005.07.688.
36. Cohen, C.J., Zhao, Y., Zheng, Z., Rosenberg, S.A., and Morgan, R.A. (2006). Enhanced antitumor activity of murine-human hybrid T-cell receptor (TCR) in human lymphocytes is associated with improved pairing and TCR/CD3 stability. Cancer Res. *66*, 8878–8886. https://doi.org/10.1158/0008-5472.CAN-06-1450.
37. Baugh, L.R., Hill, A.A., Brown, E.L., and Hunter, C.P. (2001). Quantitative analysis of mRNA amplification by in vitro transcription. Nucleic Acids Res. *29*, E29. https://doi.org/10.1093/nar/29.5.e29.
38. Rosenberg, S.A., and Restifo, N.P. (2015). Adoptive cell transfer as personalized immunotherapy for human cancer. Science *348*, 62–68. https://doi.org/10.1126/science.aaa4967.
39. Pasetto, A., Gros, A., Robbins, P.F., Deniger, D.C., Prickett, T.D., Matus-Nicodemos, R., Douek, D.C., Howie, B., Robins, H., Parkhurst, M.R., et al. (2016). Tumor- and neoantigen-reactive T-cell receptors can Be identified based on their frequency in fresh tumor. Cancer Immunol. Res. *4*, 734–743. https://doi.org/10.1158/2326-6066.CIR-16-0001.
40. Han, A., Glanville, J., Hansmann, L., and Davis, M.M. (2014). Linking T-cell receptor sequence to functional phenotype at the single-cell level. Nat. Biotechnol. *32*, 684–692. https://doi.org/10.1038/nbt.2938.
41. Lu, Y.C., Zheng, Z., Robbins, P.F., Tran, E., Prickett, T.D., Gartner, J.J., Li, Y.F., Ray, S., Franco, Z., Bliskovsky, V., et al. (2018). An efficient single-cell RNA-Seq approach to identify neoantigen-specific T cell receptors. Mol. Ther. *26*, 379–389. https://doi.org/10.1016/j.ymthe.2017.10.018.
42. McDaniel, J.R., DeKosky, B.J., Tanno, H., Ellington, A.D., and Georgiou, G. (2016). Ultra-high-throughput sequencing of the immune receptor repertoire from millions of lymphocytes. Nat. Protoc. *11*, 429–442. https://doi.org/10.1038/nprot.2016.024.
43. Turchaninova, M.A., Britanova, O.V., Bolotin, D.A., Shugay, M., Putintseva, E.V., Staroverov, D.B., Sharonov, G., Shcherbo, D., Zvyagin, I.V., Mamedov, I.Z., et al. (2013). Pairing of T-cell receptor chains via emulsion PCR. Eur. J. Immunol. *43*, 2507–2515. https://doi.org/10.1002/eji.201343453.
44. Spear, T.T., Evavold, B.D., Baker, B.M., and Nishimura, M.I. (2019). Understanding TCR affinity, antigen specificity, and cross-reactivity to improve TCR gene-modified T cells for cancer immunotherapy. Cancer Immunol. Immunother. *68*, 1881–1889. https://doi.org/10.1007/s00262-019-02401-0.
45. Aoki, H., Shichino, S., Matsushima, K., and Ueha, S. (2022). Revealing clonal responses of tumor-reactive T-cells through T cell receptor repertoire analysis. Front. Immunol. *13*, 807696. https://doi.org/10.3389/fimmu.2022.807696.
46. Mitchell, A.M., and Michels, A.W. (2020). T cell receptor sequencing in autoimmunity. J. Life Sci. *2*, 38–58. https://doi.org/10.36069/jols/20201203.
47. Wu, T.D., Madireddi, S., de Almeida, P.E., Banchereau, R., Chen, Y.J.J., Chitre, A.S., Chiang, E.Y., Iftikhar, H., O'Gorman, W.E., Au-Yeung, A., et al. (2020). Peripheral T cell expansion predicts tumour infiltration and clinical response. Nature *579*, 274–278. https://doi.org/10.1038/s41586-020-2056-8.
48. DeWolf, S., and Sykes, M. (2017). Alloimmune T cells in transplantation. J. Clin. Invest. *127*, 2473–2481. https://doi.org/10.1172/JCI90595.
49. Morris, H., DeWolf, S., Robins, H., Sprangers, B., LoCascio, S.A., Shonts, B.A., Kawai, T., Wong, W., Yang, S., Zuber, J., et al. (2015). Tracking donor-reactive T cells: evidence for clonal deletion in tolerant kidney transplant patients. Sci. Transl. Med. *7*, 272ra10. https://doi.org/10.1126/scitranslmed.3010760.
50. Barde, I., Laurenti, E., Verp, S., Wiznerowicz, M., Offner, S., Viornery, A., Galy, A., Trumpp, A., and Trono, D. (2011). Lineage- and stage-restricted lentiviral vectors for the gene therapy of chronic granulomatous disease. Gene Ther *18*, 1087–1097. https://doi.org/10.1038/gt.2011.65.
51. Gannon, P.O., Harari, A., Auger, A., Murgues, C., Zangiacomi, V., Rubin, O., Ellefsen Lavoie, K., Guillemot, L., Navarro Rodrigo, B., Nguyen-Ngoc, T., et al. (2020). Development of an optimized closed and semi-automatic protocol for Good Manufacturing Practice manufacturing of tumor-infiltrating lymphocytes in a hospital environment. Cytotherapy *22*, 780–791. https://doi.org/10.1016/j.jcyt.2020.07.011.
52. Jespersen, H., Lindberg, M.F., Donia, M., Söderberg, E.M.V., Andersen, R., Keller, U., Ny, L., Svane, I.M., Nilsson, L.M., and Nilsson, J.A. (2017). Clinical responses to adoptive T-cell transfer can be modeled in an autologous immune-humanized mouse model. Nat. Commun. *8*, 707. https://doi.org/10.1038/s41467-017-00786-z.
53. Oliveira, G., Stromhaug, K., Klaeger, S., Kula, T., Frederick, D.T., Le, P.M., Forman, J., Huang, T., Li, S., Zhang, W., et al. (2021). Phenotype, specificity and avidity of antitumour CD8(+) T cells in melanoma. Nature *596*, 119–125. https://doi.org/10.1038/s41586-021-03704-y.
54. Giordano-Attianese, G., Gainza, P., Gray-Gaillard, E., Cribioli, E., Shui, S., Kim, S., Kwak, M.J., Vollers, S., Corria Osorio, A.D.J., Reichenbach, P., et al. (2020). A computationally designed chimeric antigen receptor provides a small-molecule safety switch for T-cell therapy. Nat. Biotechnol. *38*, 426–432. https://doi.org/10.1038/s41587-019-0403-9.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Live/dead APC	Invitrogen	#L10102
CD8 FITC	Biolegend	RRID:AB 1877178, #344704
CD4 Pacific Blue	Biolegend	RRID:AB 397037, #558116, BD Biosciences
CD137 PE	Miltenyi Biotec	RRID:AB 2654986, #130-110-763
Bacterial and virus strains		
pRRL.hPGK.GP91	Barde et al.[50]	Addgene #30477
Biological samples		
PBMC from healthy donor	Buffy coats and apheresis filters from anonymous healthy donors were collected from the local transfusin center following the legal swiss guidelines.	project #P_123
TIL samples	Samples of 3 melanoma patients enrolled in a phase I clinical trial of TIL ACT were collected at baseline.	Clinical trial #NCT03475134
Chemicals, peptides, and recombinant proteins		
Actinomycin D	Sigma-Aldrich	A9415-2MG
Critical commercial assays		
MessageAmp II aRNA amplification kit	Thermo fisher scientific	AM1751
ImmunoSEQ human TCRb	Adaptive biotechnologies	
SMARTer Human TCR a/b Profiling Kit v2	Takara	#634478
Deposited data		
Raw and analyzed data	This paper	GEO: GSE225984
Experimental models: Cell lines		
293T cell line	ATCC	CRL-3216
Experimental models: Organisms/strains		
NOD.Cg-*Prkdc*^{scid} *Il2rg*^{tm1Sug} Tg(CMV-IL2) 4-2Jic/JicTac	Taconic	hIL-2 NOG mice
Oligonucleotides		
Oligos and primers used in this paper	This paper	See Tables S1–S6
Software and algorithms		
XLSTAT	Lumivero	https://www.xlstat.com/fr/
MiXCR	Mi Laboratories	https://milaboratories.com/software

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Raphael Genolet (raphael.genolet@chuv.ch).

Materials availability

This study did not generate new unique reagents.

Data and code availability

- Raw sequencing data have been deposited at GEO and are publicly available as of the date of publication. Accession number: GSE225984. Patients data will be available upon request with some restrictions.
- This paper does not report original code.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Patient samples

Leukapheresis and tumor samples were obtained from stage IV melanoma patients upon written informed consent. Peripheral blood mononuclear cells (PBMCs) were isolated from leukapheresis upon thawing and washing using the LoVo spinning membrane filtration system (Fresenius Kabi, Oberdorf, Switzerland). Autologous tumor cell lines were established from tumor fragments and maintained in RPMI 1640 supplemented with 10% fetal bovine serum (Gibco), 100IU/ml Penicillin and 100 μg/ml Streptomycin (Bioconcept). TILs were manufactured in the GMP facility at the CHUV Center of Experimental Therapeutics (CTE) in a fully GMP compliant process.[51]

For RNA stability assay, PBMC were treated with 5 μg/ml of actinomycin D (Sigma Aldrich) 30 minutes before T0. Cells were collected at T0, after 1, 2 and 4 hours. Cells were washed once with PBS before mRNA extraction.

Mice model

IL-2 NOG mice[52] (Taconic) were maintained in a Specific Pathogen Free (SPF) animal facility at the University of Lausanne under specific pathogen-free status. Six- to nine-weeks old female mice were used for the *in vivo* studies.

METHOD DETAILS

RNA extraction

For tissues or tumor biopsies total RNA was extracted using the RNeasy Mini kit according to the manufacturer's instructions (Qiagen) and resuspended in 30 μl of deionized water. Total RNA was quantified using Qubit assay (Thermo Fisher Scientific). For cells in suspension, mRNA was extracted using the Dynabeads mRNA DIRECT purification kit according to the manufacturer's instructions (Thermo Fisher Scientific) and resuspended in 30 μl of deionized water.

TCR amplification: SEQTR (SEQuencing T cell Receptor)

mRNA amplification by in vitro *transcription*

Total RNA or mRNA was amplified by *in vitro* transcription using the MessageAmp II aRNA amplification kit (Thermo Fisher Scientific) according to manufacturer's instructions, with the following modification : i) 250 ng of total RNA (tumor tissue) diluted in 10 μl or 10 μl of unquantified mRNA (cell in suspension) were used for the first strand cDNA synthesis. ii) After synthesis of the second strand, dsDNA was purified using Ampure XP beads (Beckman Coulter). 80 μl of beads were added to the dsDNA and mixed at 1800 rpm for 30''. After 5' incubation, beads were isolated on magnet until the supernatant was clear. The supernatant was then removed and the beads washed twice with the following process : with the beads still on the magnet, 200 μl of 80% EtOH were added and after 30'' incubation, EtOH was removed. After the washing steps, beads were air-dried for 2'. Once EtOH was fully evaporated, the beads were resuspended in 20 μl of deionized water and mixed at 1800 rpm for 30'. After 5' of incubation, beads were isolated on magnet until the sample was clear. dsDNA was recovered and 16 μl of it used for the *in vitro* transcription. iii) *in vitro* transcription was performed at 37°C for 14h. Synthesized complementary (c)RNA was purified with RNAclean XP (Beckman Coulter). 72 μl of beads were added to the cRNA and mixed at 1200 rpm for 30''. After incubation of 5', beads were isolated on magnet until the supernatant was clear. Supenatant was removed and the beads were then washed three times with the following process : with the beads still on the magnet, 200 μl of 70% EtOH were added and rafter 30'' incubation, EtOH was removed.After the washing steps, beads were air-dried for 2'. Once EtOH was fully evaporated, the beads were resuspended in 35 μl of deionized water and mixed at 1800 rpm for 30'. Beads were isolated on magnet until the sample was clear and cRNA collected. cRNA was then quantified using NanoDrop.

TCR ssDNA synthesis

Human single strand (ss)DNA synthesis was performed using the SuperScript III (Thermo Fisher Scientific). cRNA was diluted at concentration of 50 ng/μl. 10 μl of diluted cRNA was mixed with 1 μl of 10 mM dNTP (Promega), 1 μl of hTRAV primers mix (2 μM each) (See Table S1 for primer sequences) and 1 μl of hTRBV primers mix (2 μM each) (See Table S2 for primer sequences) and incubated at 70°C for 5' and then cool down to 50°C. After incubation, a mix containing 1 μl of 100 mM DTT, 4 μl of buffer, 1 μl of RNAsine (Promega) and 1 μl of SuperScript III was added to cRNA. The sample was incubated at 50°C for 30'' and then at 55°C for 1 hour. For the mouse repertoire, ssDNA of α and β chains were done separately. For the β chain, the same protocol than for human was used (See Table S3 for mTRBV primer sequences). For the α chain, ssDNA synthesis was performed with the SuperScript IV (Thermo Fisher Scientific). cRNA was diluted at concentration of 50 ng/μl. 10 μl of diluted cRNA was mixed with 1 μl of 10 mM dNTP (Promega), 1 μl of TRAV primers mix (2 μM each) (See Table S4 for primer sequences) and 1 μl of deionized water and incubated at 70°C for

5'. After incubation, a mix containing 1 μl of 100 mM DTT, 4 μl of buffer, 1 μl of RNAsine (Promega) and 1 μl of SuperScript IV was added to cRNA. The sample was incubated at 57°C for 30'' and then at 65°C for 1 hour. After synthesis, cRNA was digested with 2 μl of RNase DNase free (Roche) at 37°C for 40'. ssDNA was then purified using RNAclean XP beads (Beckman Coulter). 40 μl of beads were added to the cDNA and mixed at 1000 rpm for 30''. After incubation of 5', beads were isolated on magnet until the supernatant was clear. The beads were then washed three times with the following process : with the beads still on the magnet, 200 μl of 70% EtOH were added and after 30'' incubation, EtOH was removed. After the washing steps, beads were air-dried for 2'. Once EtOH was fully evaporated, the beads were resuspended in 35 μl of deionized water and mixed at 1800 rpm for 30'. Beads were isolated on magnet until the sample was clear and cDNA collected.

TCR amplification

The amplification was performed with Phusion Hot Start DNA Polymerase (NEB) and a primer in the adapter and a primer in the constant region (See Table S5 for sequences and concentration). The PCR mix was composed of 7 μl of cDNA, 1 μl of 10 μM dNTP mix (Promega), 0.4 μl of primers mix, 2 μl of buffer and 0.2 μl of polymerase. TCRs were amplified with the following PCR conditions: 98°C for 4', 20 x (98°C for 10'', 55°C for 30'', 72°C for 30''), 72°C for 2'. PCR product was purified by adding 1 μl of ExoSAP-IT (Affimetrix) and incubating at 37°C for 15' and then at 85°C for 15'. The second PCR was performed with the Phusion hot Start (NEB) to add the Illumina adapter and index. A mix containing 1 μl of 10 μM dNTP(Promega), 1 μl of NexteraXT primer index (See Table S7 for sequences and concentration), 3 μl of buffer, 0.2 μl of enzyme and 9.8 μl of deionized water was prepared and added directly to the purified PCR product. Unique dual indexes were used to reduce index hopping. The second amplification was performed with the following conditions : 98°C for 4', 25 x (98°C for 10'', 55°C for 30'' 72°C for 30''), 72°C for 2'. 5 μl of the product was run on agarose gel to verify the amplification.

TCR purification and quantification

Purification of the PCR product was performed with Ampure XP beads (Beckman Coulter). 30 μl of deionized water and 30 μl of beads were added to the PCR product and mixed at 1800 rpm for 30''. After incubation of 5', beads were isolated on magnet until the supernatant was clear. The beads were then washed twice with the following process : with the beads still on the magnet, 200 μl of 80% EtOH were added and after 30'' incubation, EtOH was removed. After the washing steps, beads were air-dried for 2'. Once EtOH was fully evaporated, the beads were resuspended in 30 μl of deionized water and mixed at 1800 rpm for 30'. After 5' of incubation, beads were isolated on magnet until the sample was clear and the TCR collected. The concentration of the purified product was then quantified using Qubit assay (Thermo Fisher Scientific).

Multiplex PCR

ImmunoSEQ

mRNA was extracted from 1 million of PBMC and resuspend in 30 μl of deionized water. 1/3 was used for SEQTR, 1/3 used for ImmunoSEQ and 1/3 used for real-time PCR. cDNA was synthesized using 3 μl of mRNA with the SuperScript III (Thermo Fisher Scientific) and random hexamer (Promega) according to the manufacturer's instructions. TCR β sequencing was performed using the ImmunoSEQ hs TCRB Kit (Adaptive Biotechnologies) according to the manufacturer's instructions. Fastq files were sand to ImmunoSEQ for data analysis.

Custom made multiplex PCR

cDNA was synthesized using 10 μl of mRNA extracted from PBMC or TILs with the SuperScript III (Thermo Fisher Scientific) and random hexamer (Promega) according to the manufacturer's instruction. RNA was digestion with RNase DNase free (Roche) for 40'. The amplification was performed with Phusion Hot Start DNA Polymerase (NEB) and a mix of primers for all TRBV segments (See Table S2) and primer for all TRAJ segments (See Table S6). The PCR mix was composed of 7μl of cDNA, 1μl of 10 μM dNTP (Promega), 0.4μl of primers mix, 2μl of buffer and 0.2μl of polymerase. TCRs were amplified with the following PCR conditions: 98°C for 4', 20 x (98°C for 10'', 55°C for 30'', 72°C for 30''), 72°C for 2'. PCR product was purified with 1 μl of ExoSAP-IT (Affimetrix) and incubating at 37°C for 15' and then at 85°C for 15'. A second PCR was performed with the Phusion hot Start (NEB) to add the Illumina adapter and index. A mix containing 1μl of 10μM dNTP(Promega), 1μl of NexteraXT primer index (See Table S7 for sequences and concentration), 3μl of buffer, 0.2μl of enzyme and 9.8μl of deionized water was prepared and added directly to the purified PCR product. The second amplification was performed with the following conditions: 98°C for 4', 25 x (98°C for 10'', 55°C for 30'' 72°C for 30''), 72°C for 2'. 5μl of the product was run on agarose gel to verify amplification. Purification of the PCR product was performed with Ampure XP beads (Beckman Coulter). 30 μl of deionized water and 30 μl of beads were added to the PCR product and mixed at 1800 rpm for 30''. After incubation of 5', beads were isolated on magnet until the supernatant was clear. The beads were then washed twice with the following process : with the beads still on the magnet, 200 μl of 80% EtOH were added and removed after 30'' incubation. After removing EtOH, beads were air-dried for 2'. Once EtOH was fully evaporated, the beads were resuspended in 30 μl of deionized water and mixed at 1800 rpm for 30'. After 5′ of incubation, beads were isolated on magnet until the sample was clear and the TCR collected. The concentration of the purified product was then quantified using Qubit assay (Thermo Fisher Scientific).

5′ RACE amplification

The 5′RACE protocol was performed with the commercially-available kit, SMARTer Human TCR a/b Profiling Kit v2 (Takara bio). mRNA was extracted from 3 millions of PBMC and resuspend in 30 μL of deionized water. $^1/_3$ was used for SMARTer, $^1/_3$ used for ImmunoSEQ and $^1/_3$ used for real-time PCR. The TCR amplification was performed according to the manufacturer's instructions.

Data were analyzed using the dedicated MiXCR pipeline (MiXCR V3.0.13, mixcr analyze takara-human-tcr-V2-cdr3) and our custom script. Similar results were obtained with both analyses.

TCR sequencing and analysis

TCR libraries were diluted at 1.8 p.m. and sequenced on MiniSeq or diluted at 60 p.m. and sequenced on NextSeq1000 instruments of Illumina with a single-end run of 150 bp. For the mouse α chain, paired-end sequencing 68-82 was performed using a custom primer (ACTGGTACACAGCAGGTTCTGGGTTCTGGATGT) in the constant region for the reverse sequence. Samples were demultiplex according to the different indexes into Fastq files. Repertoire analysis was performed with an *ad hoc* perl script that process the following way : i) identical sequences are pooled to be aligned only once. The frequency of the sequence is adjusted to correspond to the number of reads pooled. ii) 40pb of the V segment are extracted and aligned against IMGT V segments using blastn. The best match is considered only if the alignment presented less than 8 mismatches (20%) on the 40bp aligned. If two segments align with the same score, the primer region is aligned in turn against the V segments database (20% mismatch tolerated). The result of primer alignment is then compared to the result of the V segments. If a correspondence is found with one of the possible V segments, the corresponding V is attributed to the sequence. If no correspondence can be found, the sequence is considered as ambiguous, with no possibility to define the precise V segment. iii) 30pb of the J segment (upstream of the constant) are used and aligned against IMGT J segments using blastn. The best match is used only if the alignment presented less 6 mismatches (20%) out of the 30bp aligned. If two segments align with the same score, the sequence is considered as ambiguous, with no possibility to define the precise J segment. iv) Once the V and the J segments are defined, the CDR3 sequence is extracted. The CDR3 sequence is then analyzed to verify that the TCR sequence is in-frame and that no stop codon is present in the CDR3. v) Sequence with the same UMI (9 bp) are pooled together to correct for PCR and sequencing error. First the consensus CDR3 sequence is determined. To be consider as consensus, a base of the CDR3 as to be present in 75% of the reads. If no consensus sequence can be defined, the CDR3 sequences of the given UMI are not corrected. Similarly, if a CDR3 sequence present more than 2 mismatches from the consensus, the sequence is not corrected as we cannot exclude that 2 TCR sequences have the same UMI. vi) Finally, all the sequences coding for the same TCR at the protein level are pooled together to determine the abundance (number of reads) of each TCR.

For downstream analysis of the repertoire, TCR with a single read were excluded. A TCR was define as a T-cell expressing the same V segment, CDR3 sequence and J segment at the protein level. Frequencies of the TCR was calculated using only the productive TCRs, out of frame TCRs or those with a stop codon in the CDR3 were removed for analyses.

TCR metrics

Richness refers to the number of functional clonotypes identified in the samples. Shannon entropy was calculated using clonotype frequency (f) as indicated below:

$$-\sum_{i=1}^{N} f_i \, log_2(f_i)$$

while normalized Shannon entropy (i.e., evenness) was calculated by dividing the Shannon entropy by $\log_2(N)$, where N is the total number of clonotypes. Clonality was defined as 1 - normalized Shannon entropy. CL50 represents the percentage of clonotypes starting from the most frequent to the less frequent needed to reach 50% of the repertoire. To calculate homology, the sum of square differences of the frequencies was calculate for the two repertoires compared. The value was normalized by the maximum distance possible, i.e., the distance obtained if no TCRs are shared:

$$1 - \left(\frac{\sum_{i=1}^{n} (f_{ia} - f_{ib})^2}{\sum_{i=1}^{n} (f_{ia})^2 + \sum_{i=n}^{n} (f_{ib})^2} \right)$$

where f_{ia} is the frequency of the TCR obtained for the first repertoire and f_{ib} the frequency of the same TCR obtained in the second repertoire.

Single-cell transcriptome and TCR sequencing

TIL samples were thawed and cultured overnight in RPMI 1640 (Gibco) supplemented with 8% Human AB serum (Biowest), non-essential amino acids, 100mM HEPES, 1mM Sodium Pyruvate, 50μM 2-mercaptoethanol (all from Gibco), 100 IU/mL Penicillin, 100μg/mL Streptomycin (Bioconcept), 2mM L-Glutamine Solution (Bioconcept) and 3000 IU/mL IL-2 (Proleukin, Novartis Pharma Schweiz AG). TILs were resuspended in PBS + 0.04% BSA and DAPI (Invitrogen) staining was performed. Live cells were sorted with a BD FACS Melody sorter and manually counted to assess viability with Trypan blue. Cells were then resuspended at 1000 cells/μL with a viability of >90% and subjected to a 10X Chromium instrument for the single-cell analysis. The standard protocol of 10X Genomics was applied and the reagents for the Chromium Single-cell 5' Library and V(D)J library (v1.0 Chemistry) were used. 12'200 cells were loaded per sample, with the targeted cell recovery of 7'000 cells according to the protocol. Using a

microfluidic technology, single-cells were captured and lysed, mRNA was reverse transcribed to barcoded cDNA using the provided reagents (10X Genomics). 14 PCR cycles were used to amplify cDNA. Part of the material was target-enriched for TCRs and V(D)J library was obtained according to manufacturer protocol (10X Genomics). Barcoded VDJ libraries were pooled and sequenced by an Illumina HiSeq 2500 Sequencer (Rapid output, Read1: 125bp, Read2: 125bcp).

Single-cell TCR sequencing data were processed by the Cell Ranger software pipeline (version 3.1.0, 10X Genomics). For TCR expression analysis, the number of TCRα and TCRβ mRNA molecule per cell was calculated using the number of unique UMI identified in the cell. Cells with less than 3 mRNA copies of the TCR (3 UMI) were removed for the analysis. The average TCR expression for a clonotype (i.e., cells expressing the same TCR at the protein level), was calculate by dividing the total number of TCR mRNA molecules by the number of cells of the clonotype. The frequency of a TCR chain was calculated by dividing the number of mRNA copies of the given chain found in the sample by the sum of all TCRα or TCRβ mRNA copies identified in the sample.

Single cell RNAseq data were obtained from Oliveira et al.[53] From the raw gene expression matrix, we gathered the UMIs counts from every captured gene belonging to the alpha and beta TCR chains. Subsequently we computed the cell-wise sum and average of the UMI count and compared their distributions between naive and non-naive TILs subsets.

Real-time PCR

RNA was reverse transcribed using random hexamer (Promega) and SuperScriptIII (Thermo Fisher Scientific) according to the manufacturer's instructions. After the reverse transcription, RNA was digested with 2 μl of RNase DNase free (Roche) at 37°C for 40'. Real time PCR was performed with KAPA SYBR FAST (Sigma) according to the manufacturer's instructions. The forward primer was designed in the V segment and the reverse primer in the constant region. Primers were used at a final concentration of 400nM. The reactions were done in triplicate. The primer efficiency was calculated in each single reaction and used to determine the ratio of the different V in the sample.

Full length TCR cloning

mRNA used for repertoire analysis was reversed transcribed using random hexamer (Promega) and Super-Script III (Thermo Fisher Scientific). 10 μl mRNA were used for the reverse transcription with random primers (Promega) according to the manufacturer's instructions. RNA was then digested with 2 μl RNase DNase free (Roche) at 37°C for 40'. The two parts of the TCR were amplified using Phusion Hot Start DNA Polymerase (NEB) with primers in the CDR3 and the V segment or CDR3 and the constant region. A mix containing 1 μl cDNA, 1 μl dNTP(Promega), 1 μl forward primer (10 μM), 1 μl reverse primer (10 μM), 5 μl buffer, 0.2 μl polymerase and 15.8 μl deionized water was prepared and the TCRs were amplified with the following conditions: 98°C for 4', 35 x (98°C for 10'', 57°C for 30'' 72°C for 45''), 72°C for 2'. PCR products were purified using Ampure XP beads as describe above in the TCR purification and quantification section. First, fusion PCR was performed with the forward primers in the 5'end of the V and the reverse primer in the 3end of the J. The fusion PCR was performed with the Phusion Hot Start DNA Polymerase (NEB). A mix containing 1μ each purified product, 1 μL dNTP (Promega), 1 μL forward primer (10 μM), 1 μL reverse primer (10 μM), 5 μL buffer, 0.2 μL polymerase and 14.8 μL deionized water was prepared and the amplification was done with the following conditions: 98°C for 4′, 20 x (98°C for 10″, 57°C for 30″ 72°C for 45″), 72°C for 2'. PCR product were purified using Ampure XP beads as describe above in the TCR purification and quantification section. A second fusion PCR was done with the Phusion Hot Start DNA Polymerase (NEB) to combine the variable and the constant region. The forward primer was designed in the 5′end of the V segment and the reverse primer in the 3′end of the constant region. A mix containing 1μL purified product, 1 μL constant region, 1 μL dNTP (Promega), 1 μL forward primer (10 μM), 1 μL reverse primer (10 μM), 5 μL buffer, 0.2 μL polymerase and 14.8 μL deionized water was prepared and the amplification was done with the following conditions: 98°C for 4′, 20 x (98°C for 10″, 57°C for 30″ 72°C for 45″), 72°C for 2'. The final product was then purified with Ampure XP beads as describe above in the TCR purification and quantification section.

T cell electroporation and tumor cell co-culture

TCR cloning was done using PCR products containing the variable region and vectors expression the α constant, the β constant and the pCRRL-pGK lentiviral backbone. The full vector is assembled with the Golden Gate Assembly Protocol using PaqCI (NEB #R0745) and the T4 DNA Ligase (NEB #M0202), according to the manufacturer instructions. In the final construction, the 2 chains are encoded by a single open reading frame, separated by a T2A sequence. Replication-defective lentiviral particles were produced as previously described.[54]

For TCR validation, transduced cells were stained with pMHC-multimer, anti-CD8 (BD Bioscience), anti-mouse TCRβ-constant (Thermo Fisher Scientific) and Aqua viability dye (Thermo Fisher Scientific). To assess tumor reactivity, the day before the transduction, melanoma tumor cells were incubated overnight in complete medium and 200 ng/mL IFNγ (Miltenyi). The following day, tumor cells were gently detached with Accutase (Thermo Fisher Scientific) and plated at 30′000 cells/well in IFNγ pre-coated 96-well Enzyme-Linked ImmunoSpot (ELISpot) plates (Mabtech) in the presence or not of Ultra LEAF purified 20 μg/mL anti-HLA A,B,C (BioLegend). Tumor cells were incubated for 1 h at 37°C and 5% CO_2. Later on, 1 × 10^5 electroporated T cells/well were added on top of tumor cells and ELISpot plates incubated for 16–20 h in the incubator according to the manufacturer's instructions. Spots were counted on a ByoSis instrument.

Autologous tumor recognition by flow cytometry

The day before the assay, autologous tumor cells were cultured in R10 medium, RPMI 1640 GlutaMAX supplemented with 10% FBS (Gibco) and 1% Penicillin/Streptomycin (Bioconcept). Cells were detached, washed with PBS, and plated at concentration of 1.5 to 2.0 x10^6 cells/well in 24-well plates overnight in R10 medium. TILs were thawed and rested overnight in R8 medium containing IL-2 (3'000 U/ml) at a concentration of 1.5 to 2.0 x10^6 TILs/well in a 24-well plate. R8 medium consists in RPMI 1640 GlutaMAX supplemented with 8% Human AB Serum (Biowest, VWR International GmbH), 1% Penicillin/Streptomycin (Bioconcept), L-Glutamine 2 mM (Thermo Fisher Scientific), HEPES 10 mM (Thermo Fisher Scientific), 2-β-mercaptoethanol 0.05 mM (Thermo Fisher Scientific), Sodium Pyruvate 1 mM (Thermo Fisher Scientific), MEM Non-Essential Amino Acids 1/100 (Thermo Fisher Scientific). On the day of the assay, tumor cells adhered to the bottom of the plate, and R10 medium was removed by gentle aspiration. TILs were therefore co-cultured with tumor cells at an E:T ratio of 1:1 at 37°C for 6 hours. 0.5 to 1.0 x10^6 TILs were kept unstimulated as a background control. After 6 hours, cells were harvested from the plate, washed with PBS, and stained with the following antibody cocktail at the respective concentration: Live/dead APC at 1/200 (#L10102, Invitrogen, Thermo Fisher Scientific), CD4 Pacific Blue at 1/33 (RRID:AB 397037, #558116, BD Biosciences), CD8 FITC at 1/200 (RRID:AB 1877178, #344704, BioLegend), CD137 PE at 1/20 (RRID:AB 2654986, #130-110-763, Miltenyi Biotec). Cells were then washed with PBS and resuspended in PBS before acquisition on a BD LSR Fortessa flow cytometer or BD FACS Melody cell sorter (when needed).

T cell transduction and adoptive transfer in immunodeficient IL-2 NOG mice

For primary human T cell transduction, $CD8^+$ T cells were negatively selected with beads (Miltenyi) from PBMCs of a healthy donor (apheresis filter), activated and transduced as previously reported,[54] with minor modifications. Briefly, $CD8^+$ T cells were activated with anti-CD3/CD28 beads (Thermo Fisher Scientific). Lentiviral particles were added after overnight activation. Activation beads were removed after 5 days of T cell culture in R8 medium supplemented with IL-2 at 50IU/ml. At day 6, transduced T cells expressing the mouse TCRβ-constant region were sorted with a FACS ARIA III. Isolated $CD8^+$ T cells were then expanded for 10 days in R8 medium and 150U/ml IL-2 before mouse injection.

IL-2 NOG mice[52] (Taconic) were maintained in a Specific Pathogen Free (SPF) animal facility at the University of Lausanne under specific pathogen-free status. Six-to nine-weeks old female mice were anesthetized with isoflurane and subcutaneously injected with 10^6 tumor cells from melanoma patient 2. Once the tumors became palpable (at day 14), 5 × 10^6 human TCR-transduced T cells were injected intravenously in the tail vein. Tumor volumes were measured by caliper twice a week and calculated as follows: volume = length x width x width/2. Mice were sacrificed by CO_2 inhalation before the tumor volume exceeded 1000mm^3or when necrotic skin lesions were observed at the tumor site. This study was approved by the Veterinary Authority of the Canton de Vaud (under the license 3387) and performed in accordance with Swiss ethical guidelines.

QUANTIFICATION AND STATISTICAL ANALYSIS

T-tests were performed using Prism or Excel software using two tails as parameters. Pearson correlations were calculated with XLStat software. For repertoire comparison, R^2 were calculated using Spearman correlation after logarithmic transformation of the data. Statistical significance was defined as p value <0.05. Statistical test used are describe in the figure legend.