Molecular approaches to identify cancer T cell antigens and improve immunogenicity
Hos, B.J.

Citation
Hos, B. J. (2023, June 29). Molecular approaches to identify cancer T cell antigens and improve immunogenicity. Retrieved from https://hdl.handle.net/1887/3628370

Version: Publisher's Version
License: Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden
Downloaded from: https://hdl.handle.net/1887/3628370

Note: To cite this publication please use the final published version (if applicable).
Chapter 3

MC38 colorectal tumor cell lines from two different sources display substantial differences in mutanome and neo-antigen expression

Barbara Schrörs*, Brett J. Hos*, Ikra G. Yildiz, Martin Löwer, Franziska Lang, Christoph Holtsträter, Julia Becker, Mathias Vormehr, Ugur Sahin, Ferry Ossendorp, Mustafa Diken

*provided equal contributions to the manuscript

Adapted from Front Immunol. 2023;14:1102282
Abstract

The cell line MC38 is a commonly used murine model for colorectal carcinoma. It has a high mutational burden, is sensitive to immune checkpoint immunotherapy and endogenous CD8+ T cell responses against neoantigens have been reported. Here, we re-sequenced exomes and transcriptomes of MC38 cells from two different sources, namely Kerafast (originating from NCI/NIH) and the Leiden University Medical Center cell line collection, comparing the cell lines on the genomic and transcriptomic level and analyzing their recognition by CD8+ T cells with known neo-epitope specificity. The data reveal a distinct structural composition of the Kerafast and Leiden cell line genomes and different ploidies. Further, the Leiden cell line harbored about 1.3-fold more single nucleotide variations and small insertions and deletions than the Kerafast cell line. In addition, the observed mutational signatures differed; only 35.3% of the non-synonymous variants and 5.4% of the fusion gene events were shared. Transcript expression values of both cell lines correlated strongly (p = 0.919), but we found different pathways enriched in the genes that were differentially upregulated in the Leiden or Kerafast cells, respectively. Our data show that previously described neoantigens in the MC38 model such as Rpl18mut and Adpgkmut were absent in the Kerafast cell line resulting that such neoantigen-specific CD8+ T cells recognizing and killing Leiden cells did not recognize or kill Kerafast cells. This strongly indicates that at least two sub-cell lines of MC38 exists in the field and underlines the importance of meticulous tracking of investigated cell lines to obtain reproducible results, and for correct interpretation of the immunological data without artifacts. We present our analyses as a reference for researchers to select the appropriate sub-cell line for their own studies.
Introduction

Effective immunotherapy with immune checkpoint inhibitors (ICIs) correlates with the mutational burden of treated tumors.1-4 High rates of tumor-specific mutations improve the odds of MHC class I-presented mutated peptide sequences, which, due to the lack of immunologic tolerance to such neo-antigens, are more likely to be recognized by T cells as non-self. Specific T cell responses have been identified against neoantigens in cancer patients and ICIs are effective in the stimulation of neoantigen-specific responses.5-10 The relevance of this class of cancer antigens is also supported by observations that tumors are under constant immunological pressure against neoantigens, and ICIs induce a marked shift of expressed neo-epitopes.8,11-13

The identification of immunologically relevant neoantigens has become a feasible exercise due to recent technological advancements in whole-genome and -exome sequencing. These technologies are suitable for the identification of expressed non-synonymous variations (SNVs), frameshift mutations, and fusion proteins. We, and others, have successfully used this approach to identify mutation-derived epitopes in (pre-)clinical settings for the design of neoantigen-specific cancer vaccines.14-20

The MC38 adenocarcinoma colorectal cell line is a well-established and often used tumor model for the pre-clinical studies of neoantigens and immunotherapeutic approaches.13,21-26 This transplantable cell line was established in 1975 by repeated injection of the carcinogen di-methyl hydrazine in mice, and is therefore characteristic of a tumor with high mutational burden.27 Recently, this cell line was sequenced for the identification of several immunogenic neo-epitopes by Yadav and colleagues.14 Our own research identified an additional mutation in the Rpl18 gene that instigated a dominant endogenous CD8+ T cell response, while the previously identified epitope in the Adpgk gene appeared less dominant.20 Most of the mutations described by Yadav et al. we could confirm14, which was obviously the result of the same (Leiden) origin of the MC38 cell line in both studies. This cell line was in the possession of the Leiden laboratory since the mid-1990s. However, another publicly available MC38 cell line from Kerafast (NCI/NIH origin) appeared to lack expression of the published immunogenic mutations, as this line failed to activate our MC38-specific T cell lines in coculture. This raised questions about the genetic constitution and altered immunogenicity of this MC38 cell line, since the Kerafast cell line is also commonly used for immunotherapeutic studies.26,28

In this study, we re-sequenced the "Leiden" and "Kerafast" MC38 cell lines for whole-exome and transcriptomic comparison. We found major discrepancies in the mutational landscape and distinct pathways were upregulated in the Leiden or Kerafast cells which might be relevant for proposed onco-immunological studies. Several previously identified immunogenic neoantigens (i.e. mutated Rpl18 and mutated Adpgk) were lacking in the Kerafast cell line, thus only the Leiden cell line was recognized by these neoantigen-specific T cells. These findings underscore the importance of the accurate sourcing of tumor cell lines which are commonly used in the immunotherapeutic field.

Results

Comparison on genomic level

We used whole exome sequencing and RNA-seq data to investigate SNVs and indels (Supplementary Table S1), copy number alterations and gene fusions in the two MC38 cell lines from Kerafast and Leiden and found substantial differences (Figure 1A, B). While the Leiden cell line carried more SNVs and indels, the Kerafast cell line harbored more fusion genes. The overlap was 34.6%, 35.2% and 32.9% for all SNVs in exons, for all non-synonymous
SNVs in exons and for all non-synonymous SNVs in exons of expressed genes, respectively (Figure 1B). The corresponding values for indels were 24.2%, 39.1% and 37.5%. Only two of in total 37 distinct high confidence fusion gene events (5.4%) were in concordance between the cell lines. Moreover, we observed a distinct structural composition of the genomes under

Figure 1. MC38 cell lines from Kerafast and Leiden differ substantially on genomic level. (A) Circos plots showing the somatic alterations of both cell lines compared to wild type C57BL/6 mice. Outer circle: SNVs (grey) and small indels (red); second circle from the outside: CNVs, log scaled, with grey dashed lines marking copy numbers 1, 25 and 200 (Kerafast only); middle: fusion gene events. (B) Number of SNVs, indels and fusion gene events detected in Kerafast or Leiden only or shared by both cell lines. (C) Variant allele frequencies (VAF) distributions of SNVs in exons in DNA and RNA of both cell lines. VAF values of -1 indicate no coverage in RNA-seq. (D) Mutational signatures observed in both cell lines. Significance was determined with t-test followed by multiple testing correction with Benjamini-Hochberg correction.

SNVs in exons and for all non-synonymous SNVs in exons of expressed genes, respectively (Figure 1B). The corresponding values for indels were 24.2%, 39.1% and 37.5%. Only two of in total 37 distinct high confidence fusion gene events (5.4%) were in concordance between the cell lines. Moreover, we observed a distinct structural composition of the genomes under
consideration, which is indicated by a high variability of gene copy numbers (Figure 1A, middle ring of Circos plot). We determined the ploidy by matching theoretical variant allele frequency (VAF) distributions of SNVs (based on absolute copy numbers, see Methods) with the observed VAF values. This resulted in a ploidy of two for the Kerafast cell line and a ploidy of five for the Leiden cell line. The number of genes with copy number variants (CNV) included 7,516 and 26,283 genes with a reduced copy number for the Leiden and Kerafast cell lines, respectively, and 12,864 and 2,659 genes, respectively, with an increased copy number. The resultant absolute gene copy numbers showed no correlation across both cell lines (Pearson correlation coefficient -0.0031). The VAF distributions peaked at 0.25 both in the DNA and RNA data of the Kerafast cell line, while the distribution in the Leiden cell line was more heterogeneous (Figure 1C). The observed prevalence of base substitutions was mainly in concordance between the cell lines, but C>T (especially in in TCC and TCT triplets; C is the mutated base, preceded by T and followed by C or T, respectively) and T>G in CTT triplets had a higher relative abundance in the Kerafast cell line compared to other substitutions than in the Leiden cell line (Supplementary Figure S1). In the same line, we observed significant differences in the relative exposure of mutation signatures AC4 (tobacco mutagens, benzoapyrene) which had a higher relative exposure in the Leiden cell line and AC17 (unknown process) which was stronger in the Kerafast cell line (Figure 1D). Signatures AC11 (alkylating agents) and AC15 (defect DNA MMR) was found only in the Kerafast cell line and signatures AC13 (APOBEC) and AC28 (unknown process) were detected only in the Leiden cell line.

Comparison on transcriptomic level

Next, we compared the expression profiles of the two cell lines. While the normalized count data of the replicates of either cell line had a Pearson's correlation coefficient of 0.988 (Leiden) and 0.996 (Kerafast), the correlation coefficient between the cell lines was only 0.952 (Supplementary Figure S2A). The mean expression values achieved a correlation coefficient of 0.919 (Supplementary Figure S2B) and differential expression analysis between the two cell lines revealed 2,871 genes differentially upregulated in the Kerafast cell line and 9,252 genes differentially upregulated in the Leiden cell line (absolute log2foldchange > 1, adjusted p-value < 0.05; Figure 2A). The genes that were upregulated in the Leiden cell line were significantly enriched for genes involved in various KEGG (Kyoto Encyclopedia of Genes and Genomes) pathways including lysosome, glycosaminoglycan biosynthesis, ECM-receptor interaction, sphingolipid metabolism, axon guidance, mannose type O-glycan biosynthesis, and cell adhesion molecules (CAMs) (adjusted p-value < 0.05, Figure 2B). The enriched pathways were associated with different biosynthesis processes and processes regulating cell adhesion, cell-cell junction formation and cell polarity. The KEGG pathways that were significantly enriched in the genes upregulated in the Kerafast cell line were glycolysis/gluconeogenesis, pyruvate metabolism, and glutathione metabolism (Figure 2B).

Comparison on immunomic level

Despite both Leiden and Kerafast cells being of the same origin, MC38, and possessing some mutations in common, they can be distinguished based on the expression of cell-line specific mutations such as AdpgkR304M and Rpl18Q125R (Table 1). The mutations in Adpgk and Rpl18 induced endogenous CD8+ T cell responses when Leiden tumors regressed in mice treated with αPDL1 and splenocytes were expanded ex vivo upon recurrent stimulation with irradiated Leiden cells (Figure 3A) to generate antigen-specific CD8+ T cell lines. Coculture of established AdpgkR304M or Rpl18Q125R specific CD8+ T cell lines (Figure 3A) with Leiden and Kerafast cells showed a strongly reduced capacity of the T cells to recognize Kerafast cells (Figure 3B).

To further explore this difference between Leiden and Kerafast cells on the immunological level, we engineered T cells expressing TCRs against AdpgkR304M or Rpl18Q125R neo-antigens
and evaluated IFNγ secretion as well as cytotoxicity by TCR-specific T cells upon co-culture with tumor cells. Upon stimulation, Adpgk-TCR transduced T cells recognized Leiden but not Kerafast cells (Figure 3C). After co-culture with IFNγ pre-stimulated Leiden cells, Rpl18-TCR transduced T cells also showed tumor recognition (Figure 3D). IFNγ pre-stimulation of Leiden cells prior to co-culture with TCR-transduced T cells resulted in an increase (>50%) in number of IFNγ spots (Figure 3C, D). The number of IFNγ spots was comparable between Kerafast, with or without IFNγ pre-stimulation, and B16-Ova cells, our control cell line pointing out that Kerafast cell line is not recognized by T cells of AdpgkR304M or Rpl18Q125R neo-antigen specificity. Only forced expression of these neo-antigens but not Ova1257-264 in Kerafast cells via electroporation of matching neoantigen encoding RNAs resulted in significant recognition of the tumor cells by Adpgk- or Rpl18-TCR transduced T cells (Figure 3C, D).

Following tumor cell recognition via IFNγ ELISPOT, we also tested in vitro cytotoxic effects of TCR-transduced T cells on tumor cells. Adpgk- and Rpl18-TCR transduced T cells resulted in 40% and 20% lysis of Leiden cells, respectively (Figure 3E). TCR-transduced T cells caused

Figure 2. Differential expression analysis of MC38 cell lines indicates distinct transcriptomic profiles. (A) Volcano plot of the differential expression analysis between Kerafast and Leiden MC38 cells. The top 25 differentially expressed (DE) genes are labeled. (B) DE genes were subjected to pathway enrichment analysis. Significantly enriched KEGG pathways are shown (adjusted p-value < 0.05).
lysis of Kerafast cells only when these cells were forced to express the matching antigens for the TCRs. Otherwise, the percentage of lysed cells by neoantigen-specific TCRs was similar between Kerafast and B16-Ova cells.

To further explore this difference between Leiden and Kerafast cells on an immunological level, we engineered T cells expressing TCRs against AdpgkR304M or Rpl18Q125R neo-antigens and evaluated IFNγ secretion as well as cytotoxicity by TCR-specific T cells upon co-culture with tumor cells.

Upon stimulation, Adpgk-TCR transduced T cells showed recognition of Leiden but not Kerafast cells (Figure 3C). After co-culture with IFNγ pre-stimulated Leiden cells, Rpl18-TCR transduced T cells as well showed recognition (Figure 3D). IFNγ pre-stimulation of Leiden cells prior to co-culture with TCR-transduced T cells resulted in an increase (>50%) in number of IFNγ spots (Figure 3C, D). Number of counted IFNγ spots was comparable between Kerafast, with or without IFNγ pre-stimulation, and B16-Ova cells, control cell line. Forced expression of AdpgkR304M or Rpl18Q125R neo-antigens, but not OvaI257-264, in Kerafast via electroporation resulted in significant recognition of the tumor cells by Adpgk- or Rpl18-TCR transduced T cells (Figure 3C, D).

Following tumor cell recognition via IFNγ ELISPOT, we also tested in vitro cytotoxic effects of TCR-transduced T cells on tumor cells via xCELLigence. Adpgk- and Rpl18-TCR transduced T cells revealed 40% and 20% lysis of Leiden cells, respectively (Figure 3E). TCR transduced T cells caused lysis of Kerafast cells only when these cells were forced to express the matching antigens for the TCRs. Otherwise, percentage of lysed cells by neoantigen specific TCRs was similar between Kerafast and B16-Ova cells.

Table 1. Expression of previously published (candidate) neoantigens in the Kerafast and Leiden cell lines. The listed neoantigens were reported by Hos et al.20 and Yadav et al.14. The mutated amino acid in the mutated sequence is indicated in red. The variant expression calculated as variant expression = transcript expression x variant allele frequency. (K: Kerafast, L: Leiden). *: CD8+ T cell activation observed by Hos et al. and/or Yadav et al.

<table>
<thead>
<tr>
<th>gene symbol</th>
<th>mutation</th>
<th>amino acid exchange</th>
<th>mutated sequence</th>
<th>transcript expression</th>
<th>variant allele frequency</th>
<th>variant expression</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adpgk*</td>
<td>chr9:59161630</td>
<td>R304M</td>
<td>ILEELASMTNELMSSIVIQ</td>
<td>7.24</td>
<td>-</td>
<td>0.34</td>
</tr>
<tr>
<td>Cig(2)</td>
<td>chr5:134739515</td>
<td>Q196V</td>
<td>TFFRSYYPFLERLLAK</td>
<td>27.27</td>
<td>14.54</td>
<td>-</td>
</tr>
<tr>
<td>Rps11*</td>
<td>chr10:17775901</td>
<td>P45A</td>
<td>RLVEELFRAAQLANDVLOQME</td>
<td>23.54</td>
<td>19.83</td>
<td>-</td>
</tr>
<tr>
<td>Rpl18*</td>
<td>chr7:52975740</td>
<td>Q125R</td>
<td>KAGGGKIFQFALAESPK</td>
<td>1139.01</td>
<td>1100.60</td>
<td>-</td>
</tr>
<tr>
<td>Wbp11</td>
<td>chr6:136770171</td>
<td>V134L</td>
<td>QYQDAVKNQKLYEVSRIP</td>
<td>51.55</td>
<td>35.66</td>
<td>-</td>
</tr>
<tr>
<td>Adg</td>
<td>chr11:84256087</td>
<td>A506T</td>
<td>SFAFPDHITMSDDARTE</td>
<td>39.06</td>
<td>28.84</td>
<td>0.21</td>
</tr>
<tr>
<td>Cpg1</td>
<td>chr2:155903380</td>
<td>D302Y</td>
<td>GGNGDPSPSPLHLYSGTVNE</td>
<td>66.44</td>
<td>20.69</td>
<td>0.15</td>
</tr>
<tr>
<td>Dpg1*</td>
<td>chr9:44137208</td>
<td>V213L</td>
<td>EAGQKVLSTASIVFNLLELEGDRY</td>
<td>22.11</td>
<td>27.68</td>
<td>0.32</td>
</tr>
<tr>
<td>Vdg</td>
<td>chr10:82110022</td>
<td>Q269L</td>
<td>ARCAQPRALDKVVIYIKLKD</td>
<td>33.03</td>
<td>21.67</td>
<td>0.28</td>
</tr>
</tbody>
</table>
Figure 3. Comparison of Leiden and Kerafast cell lines on immunogenic level. (A) Antigen specificities of established CD8+ cell lines were analyzed by Rpl18 and Adpgk specific tetramers. (B) Established CD8+ T cells lines were analyzed for recognition of Leiden or Kerafast tumor cells by induced cytokine production after coculture with live tumor cells. IFNγ secretion by Adpgk-TCR (C) or Rpl18-TCR (D) transduced T cells upon co-culture with different tumor cells via ELISPOT assay. Data indicate mean ± SD of biological replicates (n=2). P values determined by One-way ANOVA Tukey’s multiple comparison test. (E) In vitro cytotoxic activity of Adpgk-TCR or Rpl18-TCR transduced T cells after 12h co-culture with different tumor cells Data indicate mean ± SD of biological replicates (n=3). P values were determined with respect to OTI-TCR control by One-way ANOVA Bonferroni’s multiple comparison test. n.s.: nonsignificant, *P<0.05, **P<0.01, ***P<0.001, ****P<0.0001.
Discussion

Murine tumor cell lines are a well-established tool for preclinical studies. MC38 is among the most commonly used tumor models for colorectal carcinoma and can be regarded as a “workhorse” for cancer immunotherapy research. Accordingly, MC38 is currently mentioned in more than 500 articles listed in PubMed (search term “((mc-38) OR mc38) AND tumor AND model”, 31 MAY 2022). By analyzing MC38 cells from two different sources, we revealed that there are at least two sub-cell lines. The two cell lines have a distinct genomic composition, distinct mutational signatures and share a minor portion of their non-synonymous variants (SNVs, indels) and fusions (35.3% and 5.4% respectively). This is in a similar range to that reported in a previous study in a series of human MCF7 breast cancer cell lines.48

The expression profiles of MC38 cells from Kerafast and Leiden correlated strongly, but there were still notable differences. Cell culture conditions can influence expression profiles, but the effect that we observed was very prominent with several thousands of genes being differentially upregulated in either cell line (Kerafast: 2,871 genes; Leiden: 9,252 genes). Using a reduced representation of the transcriptome that allows to infer 81% of non-measured transcripts (“L1000 assay”; see ref. 49), Ben-David and colleagues (see ref. 48) found a median of 654 genes (range: 10–1,574) that were differentially expressed by at least two-fold in pairs of MCF7 cell lines. Of note, Adpgk and Rpl18 were not differentially expressed in our analysis. Thus, both neoantigens would have the same potential to be recognized by T cells but the mutations were only present in Leiden cells. With the transfection of the neo-epitope-specific TCRs in T cells, we confirmed our findings that T cell lines raised on Leiden cells induce expansion of AdpgkR304M and Rpl18Q125R specific T cells with specificity for Leiden tumor cells while non-responsive to Kerafast cells. Induced expression of the mutated peptides by transfection rescues the recognition of Kerafast cells by the transduced T cells, thus reaffirming the lack of the mutations as the key reason for the absence of recognition of the Kerafast cells.

We further screened literature for exemplary studies addressing immunotherapeutic strategies in MC38. Yadav et al. (see ref. 14) trace back their cells to “Academisch Ziekenhuis Leiden” (or Academic Hospital Leiden, now named: Leiden University Medical Center) and the observed mutational burden is in concordance with what we found for the Leiden cells. Zhong and colleagues (see ref. 50) refer to the laboratory of Antoni Ribas at UCLA, LA, California. The sequenced ex vivo tumor material shows a mutational profile (base substitutions, mutational load) similar to our MC38 cells from Leiden. Furthermore, they find Smad4 mutated which we detected also only in Leiden cells. Other studies (e.g. ref. 51) name Kerafast as the source of their MC38 cells, but use Yadav et al. (see ref. 14) as the reference for neoantigens for their peptide vaccination. In that manuscript, the neoantigen Dpagt1mut which is present in Kerafast and Leiden cells was included in the peptide pool for vaccination. Hence, immune responses could still be observed.

Given the genetic instability and variability of tumor cell lines in general, our analyses further underline the importance of accurate tracing of tumor cell lines in the experimental design to ensure reproducible studies and avoiding artifact in data interpretation due to genomic (and thus transcriptomic as well as immunogenic) differences.
Method and materials

Animals

Female C57BL/6 Thy1.1+ donor mice were purchased from Envigo. All mice were kept in accordance with federal and state policies on animal research at BioNTech SE, Germany.

Cell lines, culture conditions and generation of viral supernatant

MC38-Leiden (Leiden) and MC38-Kerafast (Kerafast) colon carcinoma cell lines were provided by Leiden University Medical Center, Netherlands, and Kerafast, USA, respectively, and cultured under standard conditions. MC38-Leiden cells were cultured in IMDM (ATCC, 30-2005) containing 8% Fetal Bovine Serum (FBS), 2 mM L-glutamine and 50µM beta-mercaptoethanol. MC38-Kerafast cells were cultured in DMEM (ATCC, 30-2002) supplemented with 10% FBS, 10 mM HEPES and 1X nonessential amino acids (NEAA). B16-Ova melanoma cell line, ectopically expressing ovalbumin antigen, was a gift from Udo Hartwig (University Medical Center Mainz, Germany) and cultured in DMEM (Gibco) containing 10% FBS. Platinum-E cells were used for generation of MLV-E pseudotyped viral particles for different TCRs and maintained under standard conditions in DMEM (Gibco) supplemented with 10% FBS. The cells were transfected with TransIT-LT1 (Mirus) based on manufacturer’s instructions. Retroviral supernatants were collected 48 and 72 h after transfection. The titers were determined using mCAT cells as described in ref. 29.

High-throughput sequencing and read alignment

Exome capture from MC38 cell lines and C57BL/6 mice were sequenced in duplicate using the Agilent Sure Select Kit and Agilent SureSelectXT Mouse All Exon exome capture assay. Oligo(dT)-isolated RNA for gene expression profiling of the MC38 cell lines was prepared in duplicate with Illumina’s TruSeq stranded Library Prep Kit. Libraries were sequenced on an Illumina HiSeq2500 or NovaSeq6000 (2×50 nt). DNA-derived sequence reads were aligned to the mm9 genome using bwa (ref. 30; default options, 0.7.10). RNA-derived sequence reads were aligned to the mm9 genome using STAR (ref. 31; default options, version 2.1.4a). The sequencing reads are available in the European Nucleotide Archive (see Data Availability Statement).

Mutation detection

Strelka2 (ref. 32; default options for whole exome sequencing, version 2.9.9) was used to call somatic SNV and short insertion/deletion (indel) on each cell line or normal library replicate pair individually.

DNA copy number calling

Absolute copy numbers were called from exome capture data as described before (see ref. 33) using Control-FREEC (ref. 34; version 11.5).

Mutation signatures

Mutation signatures (see ref. 35) were computed with the R package YAPSA (ref. 36; default settings, version 1.10.0).
Fusion gene detection

Fusion genes were detected with EasyFuse (version 1.3.6) using a “wisdom of crowds” approach as detailed before (see ref. 37). Entries in the “references” and “other_files” sections of the EasyFuse configuration were changed to Ensembl GRCm38.95. Data for both MC38 cell lines was available in two replicates. Intersection of fusion gene events (i.e. unique breakpoint IDs (BPID)) from both replicates with a prediction probability score ≥ 0.5 was taken from each origin to obtain a high confidence dataset. Fusion events reported in chrY were not considered.

Circos plots

Somatic alterations in each cell line (SNVs, INDELs, fusion genes and copy number variations) were visualized in circos plots with R package Circlize (ref. 38; version 0.4.11). Genomic coordinates of the fusion event breakpoints were converted to mm9 with liftOver (see ref. 39). Breakpoint 1 of the fusion event with BPID “X:170018795:+_X:169984999:+” could not be converted. For the visualization, it was manually set to X:166456727 at the same genomic distance to breakpoint 2 (X:166422931) in mm9.

Transcriptome profiling

Transcript abundance estimation was done with kallisto (see ref. 40; default options, version 0.42.4) on each cell line library replicate individually using the mean transcripts per million (TPM) per transcript final value. Differential expression analysis was performed using DESeq2 (see ref. 41; version 1.24.0) with Leiden cell line as “control” and the transcript counts reported by kallisto, summarized by adding up the counts of the respective transcripts associated with each gene. Enriched pathways (KEGG 2019 Mouse) in differentially up- or downregulated genes were determined using Enrichr (see ref. 42).

Construction of T cell receptor vectors

The codon-optimized and synthesized individual TCR-alpha and TCR-beta sequences reactive against AdpgkR304M, Rpl18Q125R and Ova257-264 antigens (Eurofins Genomics) were cloned into the retroviral vector MP71 for stable expression in murine T cells. TCR genes were connected to firefly luciferase and eGFP reporter genes by 2A-splice elements (see ref. 43).

Retroviral engineering of murine T cells

Splenocytes of naïve C57BL/6-Thy1.1+ mice were pre-activated by 2 mM/mL Concanavalin A (ConA) (Sigma) in T cell media, RPMI1640-GlutaMAX supplemented with 10% FBS, 1x NEAA, 1 mM sodium pyruvate, 10 mM HEPES, 50 μM β-Mercaptoethanol, 50 IU/mL Penicillin and 50 μg/mL Streptomycin, in the presence of 450 IU/mL rh IL-7 and 50 IU/mL rh IL-15. 24 h after activation, cells were gently spun down (1h, 37°C, 300 x g) and incubated on MLV-E-pseudotyped gamma-retroviral vector pre-coated-RetroNectin-plates (Takara). After additional overnight cultivation, spin-down transduction was repeated on freshly viral particles coated plates. 72 h after initial pre-activation, ConA was removed from culture and lymphocyte layer was isolated by Ficoll-Hypaque (Amersham Biosciences) density gradient centrifugation. Non-transduced T cells used as control for some experiments underwent the same ConA-activation procedure. Transgene expression on transduced murine T cells were measured via flow cytometry.

RNA constructs and in vitro transcription

Plasmid templates for in vitro transcription of antigen-encoding RNAs, i.e. Adpgk-RNA and
Rpl18-RNA, were based on pSTI vector. They were designed to encode 27 amino acids with the mutated amino acid at the central position (position 14). As a control, Ova-RNA encoding for Ova257-264 (SIINFEKL) amino acid as well as enhanced green fluorescent protein (eGFP) was employed (see ref. 44). In vitro transcription and capping with β-S-anti-reverse cap analog (ARCA) was performed as described in (see ref. 44).

**Electroporation of target cells**

MC38-Leiden and –Kerafast cells were resuspended in X-VIVO 15 (Lonza) and electroporated in 4-mm cuvettes (Bio-Rad) with an ECM 830 Square Wave Electroporation System (BTX) (300V, 15 ms, 1 pulse) after addition of 2 µg antigen encoding RNA. The cells were co-electroporated with 2 µg eGFP RNA as an electroporation control. Cells were diluted immediately in culture medium directly after electroporation. 16-20 h post electroporation, cells were harvested to be used in the downstream applications such as IFNγ ELISPOT or cytotoxicity assay. The transfection efficiency was assessed based on GFP expression via flow cytometry.

**Flow cytometry**

Transduction efficiency and TCR expression by T cells following transduction was measured via flow cytometry. The monoclonal antibodies against mouse CD8α-PE-Cy7 (BioLegend; clone:53-67), CD8α-PE-Cy7 (ThermoFisher; clone:5H10), and CD8α-APC-R700 (BD Biosciences; clone 53-67) were used. Cytokine production by T cell lines was analyzed with TNFα-MP6-XT22 (BioLegend; clone MP6-XT22) and IFNγ-PE-Cy7 (BD Biosciences; clone XMG1.2) antibodies. TCR expression after transduction was evaluated based on tetramer staining. The following tetramers were used; Adpgk-tetramer-APC (ASMTNMELM-H-2-Db), Adpgk-tetramer-PE (ASMTNMELM-H-2-Db), Rpl18-tetramer-APC (KILTFDRL-H-2-Kb) and Oval-tetramer-APC (SIINFEKL-H-2-Kb) (all MBL). TCR transduced T cells were stained for 30 min at 4°C. PBS containing 5% FBS and 5 mM EDTA was used as washing and staining buffer. Acquisition and analysis were performed on a BD FACS Cantoll and FlowJo software, respectively.

**MC38-Leiden-specific T cells**

T cell lines originate from anti-PDL1 (clone MIH-5) treated, MC38 immune mice as described by Sow *et al.* (see ref. 46), and ex vivo established via coculture of splenocytes with irradiated MC38 in IL2 supplemented (5 Cetus Units) medium (as described by Hos *et al.*, see ref. 20). Recognition of live MC38 cells (Leiden and Kerafast) by T cell lines was determined by cytokine production after o/n coculture in a 5:1 (effector : target) ratio and 2 µg/mL brefeldin A (Sigma-Aldrich) by flow cytometry.

**IFNγ ELISPOT**

Adpgk-, Rpl18- or OTI-TCR-transduced T cells were cultured overnight at 37°C on anti-IFNγ (Mabtech, clone AN18) pre-coated Multiscreen filter plates (Merck Millipore). 1x10^5 transduced T cells were stimulated with 5x10^4 tumor cells, i.e. B16-Ova, MC38-Leiden or –Kerafast cells (untreated or pre-treated overnight with 20 ng/mL IFNγ), or MC38-Kerafast cells electroporated with Adpgk-, Rpl18- or Oval-RNA. The spots were visualized with a biotin-conjugated anti-IFNγ antibody (Mabtech) followed by incubation with ExtrAvidin-Alkaline Phosphatase (Sigma-Aldrich) and BCIP/NBT substrate (Sigma-Aldrich). Plates were scanned using CTL's ImmunoSpot® Series S five Versa ELISpot Analyzer (SSVersa-02-9038) and analyzed by ImmunoCapture V6.3 software. The samples were tested in duplicates and spot counts were summarized as means of technical duplicates.
Cytotoxicity assay

TCR mediated cytotoxicity was evaluated using the xCELLigence system (OMNI Life Science). Cell index (CI) impedance measurements were performed according to manufacturer’s instructions. Target cells MC38-Leiden and –Kerafast were seeded at a concentration of 4x10^4 and 2x10^4 cells per well, respectively, in E-plate 96 (ACES Biosciences Inc.). After 20-24 h, TCR transduced murine T cells were added at 60:1 E:T (effector:target) ratio onto tumor cells in a final volume of 200 µL and monitored every 30 min for 72 h by xCELLigence device. The maximum CI corresponds to the minimal lysis (L_min), tumor cells incubated with irrelevant TCR (OTI-TCR) transduced T cells. The minimum CI corresponds to the maximum lysis, tumor cells co-incubated with 2 mM Staurosporine (Sigma) in the absence of any T cells. Percent lysis, after 12h co-incubation for each sample, was calculated using the following equation, . Then, the specific lysis for each neoTCR was calculated by normalizing the % Lysis_{NeoTCR} to % Lysis_{Staurosporin} (positive control, 100% lysis).

Statistical analysis and depiction of data

All results are represented with +/- SD of technical duplicates or triplicates. Statistical analysis for each experiment is described in the corresponding figure legend. All statistical analyses were performed using GraphPad PRISM 9 or R Studio (see ref. 47).
References


Supplemental figures

**Figure S1.** Abundance of nucleotide substitutions in both cell lines with respect to nucleotide triplets. Only combinations with a relative abundance > 0.01 in at least one of the cell lines are shown.

**Figure S2.** Correlation of expression profiles. (A) Pearson’s correlation coefficients of normalized count data between sequencing replicates and between MC38-K and MC38-L cell lines. (B) Scatter plot of log-transformed FPKM expression values (mean of replicates) of MC38-K and Leiden cells. The red line MC38-L the linear regression (correlation coefficient: 0.919).