nature biotechnology

Article

https://doi.org/10.1038/s41587-024-02161-y

Prediction of tumor-reactive T cell receptors from scRNA-seq data for personalized T cell therapy

Received: 13 June 2023

Accepted: 1 February 2024

Published online: 7 March 2024

Check for updates

C. L. Tan[1,2,3,4], K. Lindner[1,2,3,5], T. Boschert[1,2,3,4,6], Z. Meng[7,8,9], A. Rodriguez Ehrenfried[4,6,8], A. De Roia[4,10], G. Haltenhof[1,3], A. Faenza[11], F. Imperatore[11], L. Bunse[1,2,3], J. M. Lindner[12], R. P. Harbottle[10], M. Ratliff[13], R. Offringa[7,8,9], I. Poschke[1,2,5], M. Platten[1,2,3,5,6,14,15] ✉ & E. W. Green[1,2,3,15] ✉

The identification of patient-derived, tumor-reactive T cell receptors (TCRs) as a basis for personalized transgenic T cell therapies remains a time- and cost-intensive endeavor. Current approaches to identify tumor-reactive TCRs analyze tumor mutations to predict T cell activating (neo)antigens and use these to either enrich tumor infiltrating lymphocyte (TIL) cultures or validate individual TCRs for transgenic autologous therapies. Here we combined high-throughput TCR cloning and reactivity validation to train predicTCR, a machine learning classifier that identifies individual tumor-reactive TILs in an antigen-agnostic manner based on single-TIL RNA sequencing. PredicTCR identifies tumor-reactive TCRs in TILs from diverse cancers better than previous gene set enrichment-based approaches, increasing specificity and sensitivity (geometric mean) from 0.38 to 0.74. By predicting tumor-reactive TCRs in a matter of days, TCR clonotypes can be prioritized to accelerate the manufacture of personalized T cell therapies.

The success of tumor infiltrating lymphocyte (TIL) therapy trials in metastatic melanoma shows that TILs contain a fraction of tumor-reactive T cells that can be harnessed for adoptive cell therapy[1]. This success is more limited in non-melanoma cancer types[2] where the baseline fraction of experimentally verifiable, tumor-reactive CD8$^+$ T cells is low—often not exceeding 0.5% (ref. 3). While the fraction of tumor-reactive T cells can be enriched before reinfusion via cell expansion, this process can exhaust the T cells, compromising their tumor-killing efficacy[4] and leading to clonal depletion[5]. In contrast, personalized transgenic T cell therapies seek to identify and reinfuse defined tumor-reactive T cell receptors (TCRs), either in patient-autologous T cells[6] or in induced pluripotent stem cell-derived, hypoimmunogenic (allogeneic) T cells[7]. While this generates a highly efficacious product, identifying tumor-reactive TCRs is a 'needle in a haystack' problem[8].

[1]CCU Neuroimmunology and Brain Tumor Immunology, German Cancer Research Center, Heidelberg, Germany. [2]German Cancer Consortium, Core Center Heidelberg, Heidelberg, Germany. [3]Department of Neurology, Medical Faculty Mannheim, Mannheim Center for Translational Neuroscience, Heidelberg University, Mannheim, Germany. [4]Faculty of Biosciences, Heidelberg University, Heidelberg, Germany. [5]Immune Monitoring Unit, National Center for Tumor Diseases, Heidelberg, Germany. [6]Helmholtz Institute for Translational Oncology, Mainz, Germany. [7]Department of General, Visceral and Transplantation Surgery, University Hospital Heidelberg, Heidelberg, Germany. [8]Division of Molecular Oncology of Gastrointestinal Tumors, German Cancer Research Center, Heidelberg, Germany. [9]Sino-German Laboratory of Personalized Medicine for Pancreatic Cancer, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China. [10]DNA Vector Laboratory, German Cancer Research Center, Heidelberg, Germany. [11]Cellply Srl, Bologna, Italy. [12]BioMed X GmbH, Heidelberg, Germany. [13]Department of Neurosurgery, University Hospital Mannheim, Mannheim, Germany. [14]German Cancer Research Center–Hector Cancer Institute at the Medical Faculty Mannheim, University of Heidelberg, Mannheim, Germany. [15]These authors contributed equally: M. Platten and E. W. Green. ✉e-mail: m.platten@dkfz.de; e.green@dkfz.de

https://doi.org/10.1038/s41587-024-02161-y

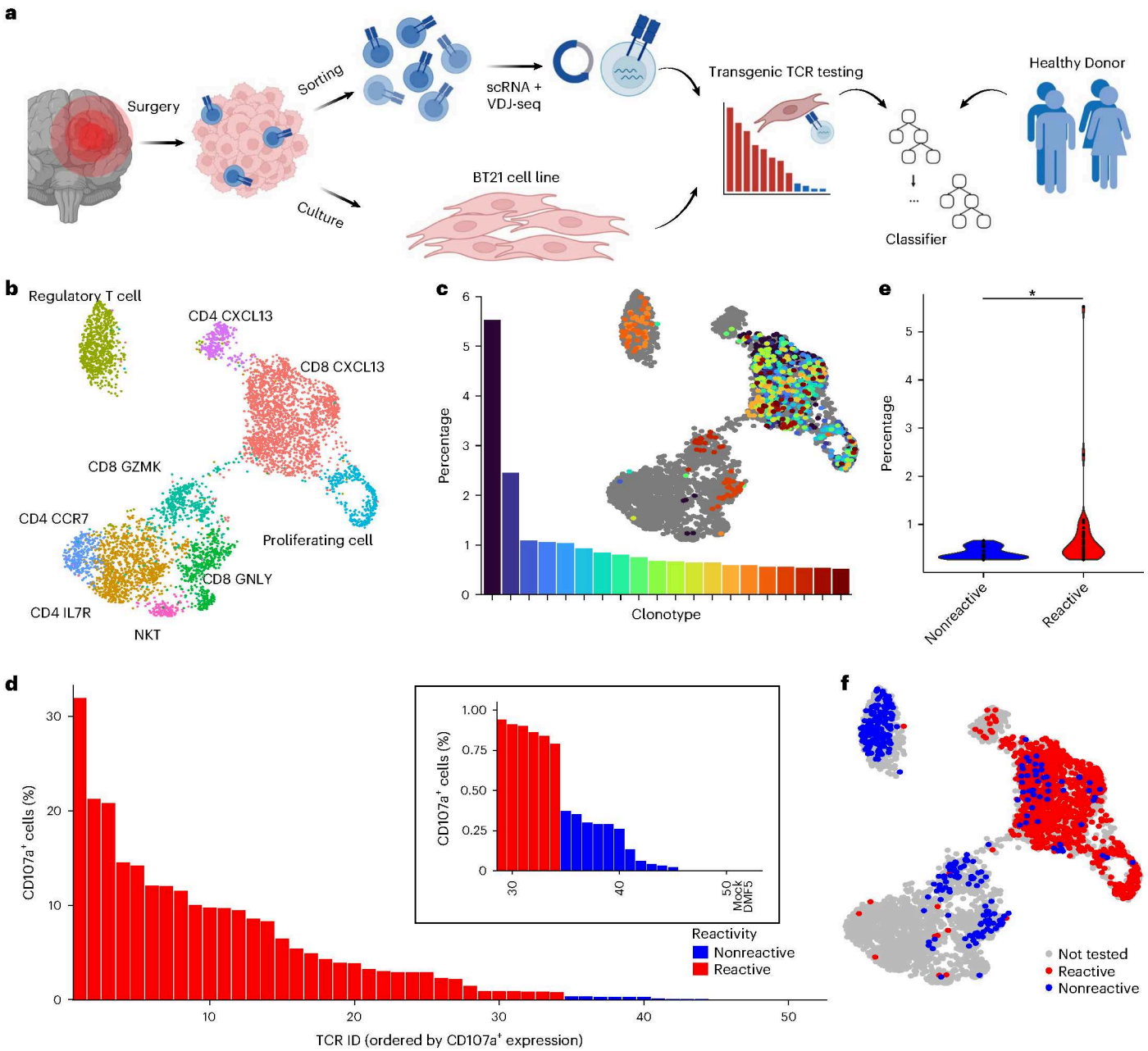


Fig. 1 | BT21 cell line accurately models resected metastatic lesion, allowing high-confidence experimental TCR tumor-reactivity testing. a, An overview of the experimental and computational pipeline underlying the predicTCR classifier: TILs are sorted and subject to scRNA + VDJ-seq, while adjacent resected tumor material is used to establish the BT21 tumor cell line. TCR reactivity data are then integrated with scRNA + VDJ-seq data to train the predicTCR classifier, which is later tested on externally generated TIL datasets from diverse tumor types. **b**, Unsupervised clustering (UMAP plot) of scRNA-seq data of TILs ($n = 5{,}651$) recovered from brain metastasis sample, with key T cell subtypes annotated. **c**, The percentage frequency of the top 20 TIL TCR clonotypes and their distribution projected onto the UMAP, showing that cells of the same clonotype can occupy diverse phenotypic states. **d**, T cells transfected with one of the 50 most frequently occurring TIL-derived TCR clonotypes (representing 58 distinct TCR α/ß chain pairs) are cocultured with BT21 cells; the resulting levels of CD107a (as quantified by flow cytometry, gated on mTCRβ$^+$ cells, which express the transgenic TCR as a chimera with the murine constant domain) demonstrate whether a given TCR clonotype recognizes the BT21 cell line. For details of settings per TCR reactivity threshold, see Methods. DMF5 is the HLA mismatched negative control TCR. **e**, BT21-reactive TCR clonotypes are more frequent than nonreactive clonotypes in the TIL population. **f**, BT21 reactivity testing results projected onto the UMAP plot (**b**).

Current techniques place emphasis on tumor antigens, using mutanome analysis to determine the most likely immunogenic neoepitopes to be screened experimentally against TCRs recovered from TILs[9]. This is a technically challenging and time-consuming endeavor: only a fraction of predicted neoepitopes represent physiologically relevant, naturally processed T cell epitopes. Furthermore, while substantial focus has been placed on tumor-specific, single-nucleotide variant (SNV)-derived neoantigens as the source of TCR epitopes, this neglects antigens generated through diverse mechanisms that are only recently beginning to be understood. These include complex mutations such as frame shifts, gene fusions and aberrant gene splicing, as well as novel targets arising though transposable element activation[10], cell stress-induced tryptophan bumps[11], aberrant posttranslational modifications[12], unannotated open reading frames[13] and even from

https://doi.org/10.1038/s41587-024-02161-y

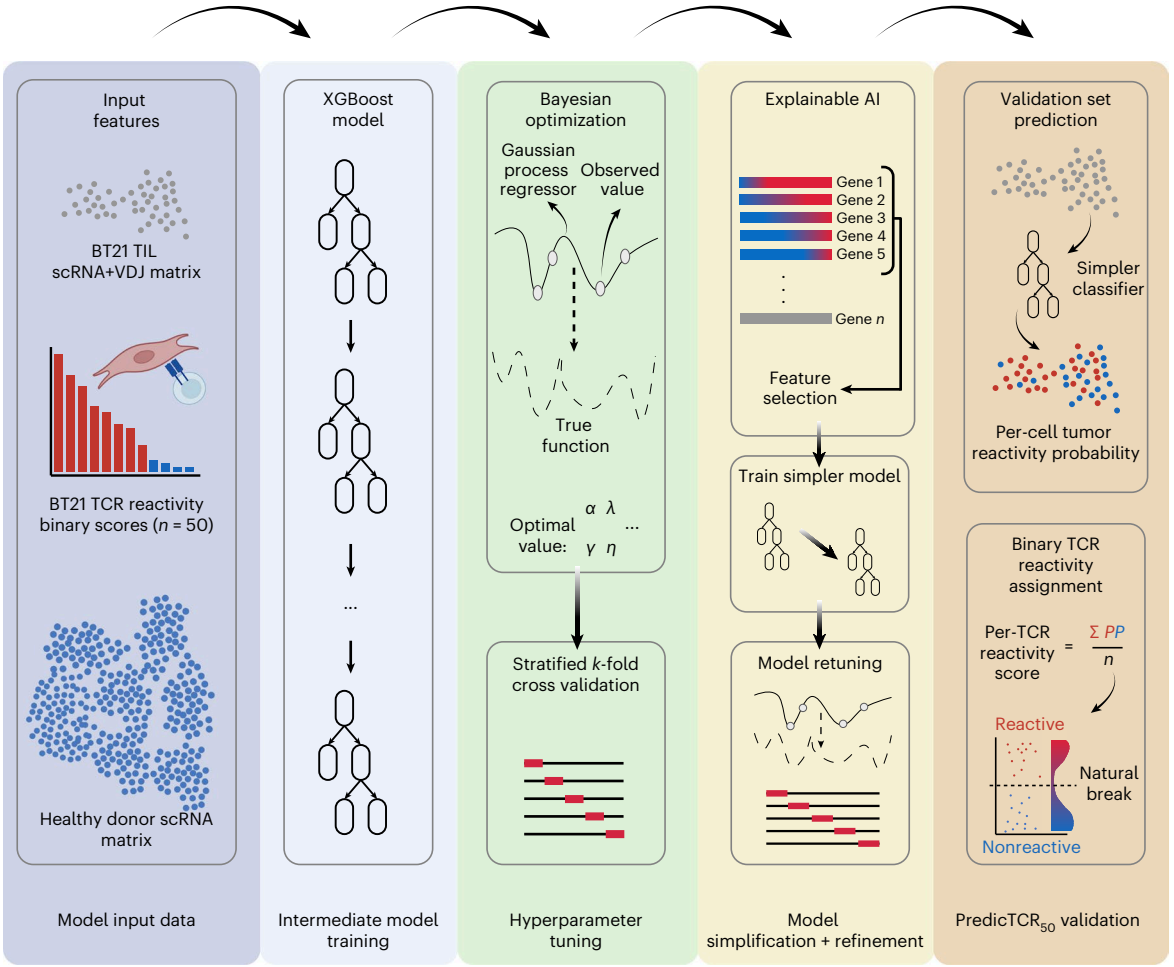


Fig. 2 | PredicTCR$_{50}$ classifier training strategy. ScRNA data from healthy donors, as well as scRNA + VDJ and experimentally derived tumor-reactivity data for the 50 most frequent TCR TIL clonotypes from sample BT21, were used to train an intermediate model using XGBoost. Due to the sparse nature of scRNA data, we optimized this intermediate model by first performing Bayesian optimization to tune hyperparameters with stratified *k*-fold cross-validation. Subsequently we identified the top features (that is, genes) in this intermediate model using explainable AI SHAP, and then trained a simpler model using only these features to prevent overfitting to the training data. This simpler model was retuned as before and then applied to the remaining BT21 TIL data. Per-cell reactivity probabilities calculated by the classifier were averaged for each TCR clonotype, and the Fisher–Jenk natural break was used to determine the appropriate minimum threshold for calling TCRs as tumor reactive.

intracellular pathogens[14]. Together, these ensure that a tumor-focused, antigen-centric approach is both slow and inefficient in identifying suitable tumor-reactive TCRs for use in personalized therapies, thus raising costs and limiting clinical application.

We hypothesized that the identification of tumor-reactive TCRs could be accelerated by developing a TCR-centric, antigen-agnostic approach: ascertaining TCR sequence and tumor reactivity directly from T cells using single-cell combined RNA + VDJ sequencing (scRNA + VDJ-seq). We have previously shown that tumor-infiltrating T cells expressing a TCR against a tumor-specific neoepitope in a vaccinated patient with glioma could be distinguished from bystander T cells on the basis of their expression of *CXCL13* and *CD40LG*[15]. This observation has been extended by other groups using cluster-based differential gene expression analyses to generate multigene 'signatures' of tumor-reactive TILs in melanoma[16–18], lung cancer[19,20], gastrointestinal cancer[21], pancreatic ductal adenocarcinoma (PDAC)[22] and metastatic cancer[23]. The reported gene signatures are only partially overlapping, implying that there may be tumor type-specific transcriptional features in TILs.

We postulated that molecular events in the process of T cell activation upon recognition of a tumor antigen are specific for tumor antigens and independent of tumor type. While differences in published signatures of T cell activation might reflect bona fide differences (for example, as a result of distinct tumor microenvironments), they might also reflect genes playing nonessential roles in T cell activation. In addition to this, the process of validating the tumor reactivity of a TCR requires the generation of tumor models that accurately recapitulate the mutational landscape and epitope processing capacity of the tumor—a process complicated by the spatial heterogeneity of many tumors. The consequence of this is that existing datasets might be noisy due to false negative TCR testing results, in which the tumor model lacks many target epitopes found in the primary tumor. Furthermore, the cost-intensive nature of these experiments and the desire to discover therapeutically useful TCRs has meant that experiments have typically focused on validating TCR clonotypes most likely to be tumor reactive rather than unbiased TCR cloning. This bias may complicate the identification of confounding transcriptional signatures not essential for T cell activation in existing data.

We reasoned that resolving these issues would allow tumor-reactive TILs to be identified regardless of tumor type from single-cell RNA sequencing (scRNA-seq) data alone. Furthermore, by cloning TCRs in an unbiased fashion and including large amounts of negative training data, a machine learning classifier could be trained to identify tumor-reactive TCR clonotypes from scRNA + VDJ-seq data in an automated manner.

Deep screening identifies tumor-reactive TCR from TILs

In this study, we set out to identify a tumor sample from which we could sequence TILs and derive a tumor cell line that appropriately recapitulated the primary tumor to allow for high-confidence TCR tumor-reactivity testing to generate a classifier training dataset (Fig. 1a). As sequencing is destructive, adjacent tumor pieces must be used for tumor and TIL sequencing as well as tumor cell line establishment. Tumor mutational heterogeneity (which correlates with TIL heterogeneity[24]) leads to TILs from one tumor piece recognizing antigens absent from the tumor cell line generated from a distal piece, resulting in false negatives during TCR testing that lower the quality of the training dataset. We therefore chose to use a metastatic tumor, as monoclonal metastasis seeding events represent genetic bottlenecks that fix mutations[25], maximizing the similarity between primary tumor and resultant cell line. We further hypothesized that a metastasis derived from the brain—which has a degree of immune privilege—might result in improved phenotypic separation between bystander and infiltrating tumor-reactive T cells. We identified a metastatic brain tumor from a 62-year-old male patient previously diagnosed with melanoma, which was established as a tumor cell line hereafter termed BT21. Whole-exome sequencing showed that BT21 was a suitable model of the metastatic tumor, sharing 245 of the 268 functional SNVs (Extended Data Fig. 1 and Source data), and constitutively expressing major histocompatibility class I (MHC I) complexes required for epitope presentation and TCR testing (Extended Data Fig. 2).

Unsupervised clustering of scRNA-seq of TIL-derived T cells (n = 5,651, hereafter referred to as TILs) showed the presence of distinct clusters expressing known markers of T cell activation including *CXCL13*, *GZMK* and *GNLY* (Fig. 1b). Single-cell VDJ sequencing (scVDJ-seq) of TILs showed the presence of expanded TCR clonotypes, with one clonotype representing over 5% of all clones—a signal indicative of tumor reactivity due to local T cell expansion (Fig. 1c). Additionally, TCR clonotypes found in the scRNA + VDJ of TILs could also be identified in the RNA-seq data derived from a distinct piece of tumor tissue, suggesting that the source tumor was relatively homogeneous in terms of T cell infiltration and presumably the underlying mutational landscape (Supplementary Table 1).

We cloned the most frequently occurring α/ß TCR chain pairs (n = 58) from the TIL population (representing 50 distinct TCR clonotypes as some T cells express two productive α chains). TCRs were transfected into expanded healthy donor peripheral blood mononuclear cells (PBMCs) and screened for reactivity against the BT21 cell line using a flow cytometry-mediated readout of T cell activation ($CD107a^+$) corrected for per TCR background tonic signaling. A conservative threshold was set to determine TCR reactivity (Methods and Extended Data Fig. 3a,b). We found 34/50 TCRs to be tumor reactive (Fig. 1d and Source data), and showed that there was no significant difference in transgenic TCR expression between reactive and nonreactive TCRs (Extended Data Fig. 3c). Tumor-reactive TCR clonotypes were significantly more expanded in the TIL population than nonreactive clonotypes (Fig. 1e) and individual cells expressing tumor reactive TCRs could occupy various states (Fig. 1b,c,f).

Development of predicTCR$_{50}$ classifier from scRNA + VDJ data

Using the TCR reactivity dataset we established for BT21, we set out to build a machine learning classifier that could accurately and robustly predict tumor reactivity of TIL-derived TCRs based on scRNA + VDJ-seq data using the strategy illustrated in Fig. 2. We first converted the 50 experimentally determined tumor reactivities into a binary label for each TIL cell expressing a tested TCR clonotype and used the corresponding gene expression matrix for those cells as input to train machine learning frameworks. The predictive performance of several machine learning frameworks was evaluated using the area under the receiver operating characteristic (ROC) curve (AUC) metric that ranges between 1 (perfectly predictive), 0.5 (no discrimination capacity between groups) and 0 (reciprocating classes). When making a classifier, a threshold must be set to discriminate between reactive and nonreactive states; the AUC metric assesses the best possible performance of a model on a given dataset by varying the threshold value. This preliminary comparison found eXtreme Gradient Boost (XGBoost)[26] to be the most suitable framework (Extended Data Fig. 4). While XGBoost performs particularly well due to its ability to update subsequent decision trees during boosting, it has additional advantages for analysis as it effectively implements within-tree parallelization and is able to handle dropout data commonly found in scRNA datasets. Importantly, XGBoost also incorporates regularization to prevent overfitting, which otherwise limits the ability to generalize a model to new datasets.

We added scRNA data from ten healthy donor PBMC samples generated by three independent groups; this produced a maximally diverse negative control dataset for training. Altogether, a total of 112,960 cells were used for training, of which 1,461 cells were TILs from BT21; the imbalanced nature of the training data required careful optimization of data weighting (Methods). XGBoost hyperparameters were tuned using stratified k-fold cross-validation with Bayesian optimization, using 70% of the TCRs for training and 30% for testing. To reduce the complexity of our model—important to prevent overfitting that would limit the performance of the classifier on new samples—we identified the key features (that is, genes) determining model performance using explainable artificial intelligence (AI) SHapley Additive exPlanations (SHAP)[27]. We then repeated hyperparameter optimization using only these features.

The probability of tumor reactivity was calculated for each individual T cell using the model, and a mean score then calculated for each TCR clonotype using the data from scVDJ-seq (as TILs expressing a given TCR clonotype may be found occupying various phenotypic states from naive to exhausted; Fig. 1b,c). Finally, the minimum reactivity score required for a TCR clonotype to be called as being tumor reactive was calculated using Fisher–Jenk break optimization, a deterministic statistical analysis that can set sample-specific thresholds. We named the resulting classifier 'predicTCR$_{50}$'.

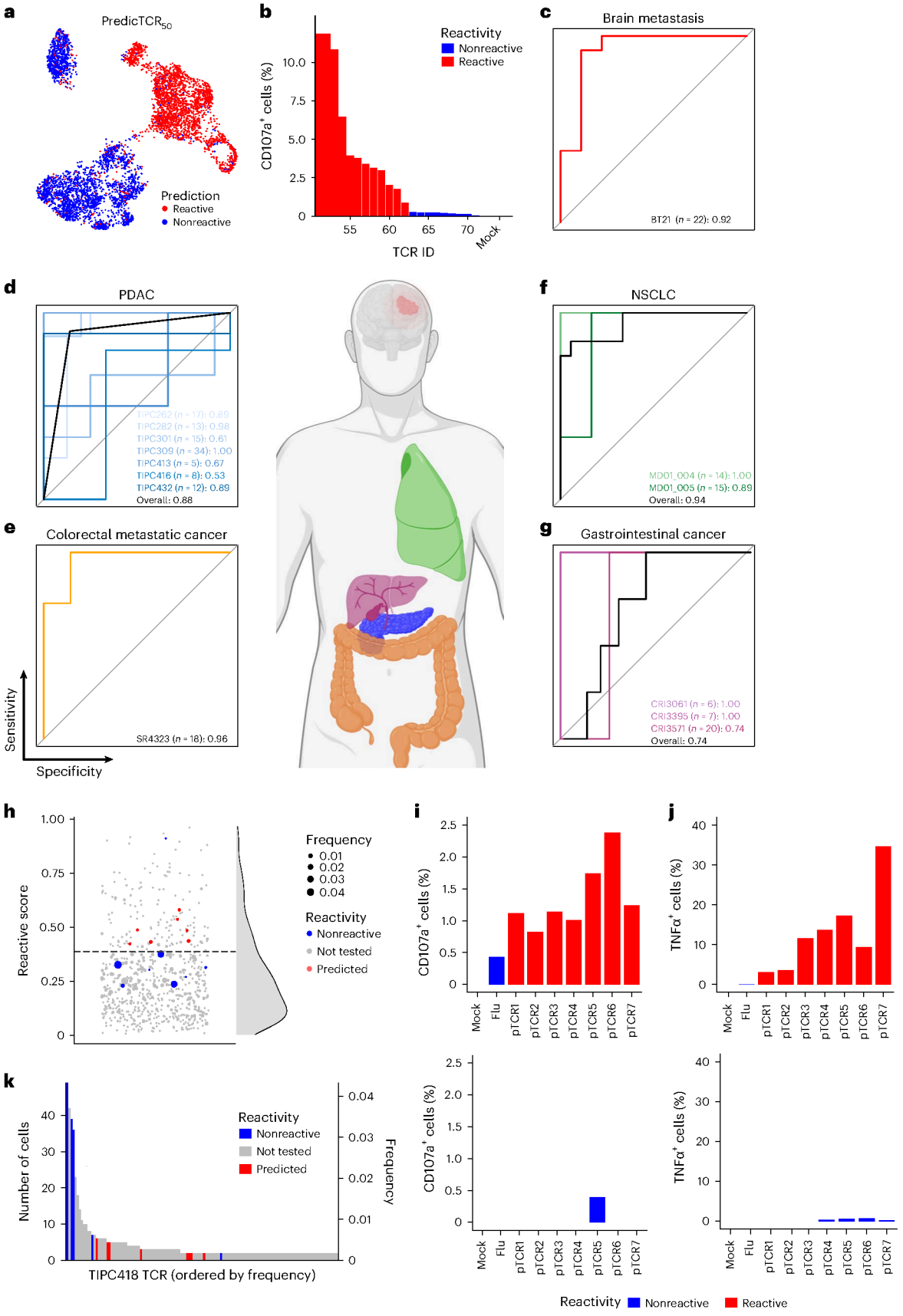
PredicTCR$_{50}$ prediction performance in brain metastasis

We used predicTCR$_{50}$ to generate tumor-reactivity predictions for all TILs recovered from the BT21 metastasis, with the per-clonotype

Fig. 3 | PredicTCR accurately predicts tumor-reactive TCRs in diverse tumor types. a, A UMAP plot as in Fig. 1 overlaid with predicTCR$_{50}$ per-cell tumor-reactivity predictions. **b**, An additional 22 TCR clonotypes (29 distinct TCR α/ß chain pairs) were tested for reactivity against the BT21 cell line. **c**, The performance of predicTCR$_{50}$ in prospective prediction of tumor-reactive TCR in BT21 patient. **d**–**g**, The performance of predicTCR in predicting TCR tumor reactivity in published scRNA + VDJ datasets with TCR reactivity data available: seven PDAC samples from Meng et al.[22] (**d**), one colon metastasis[23] (**e**), two NSCLC[19] (**f**) and three gastrointestinal cancers[21] (**g**). The metrics were calculated by clonotype, with the number of TCR clonotypes for each sample and the AUC value listed. The overall performance was assessed using all available TCRs per cancer modality. Additional metrics and details of the sequencing technology and reactivity testing method used for each sample are listed in Table 3. **h**, PredicTCR reactivity predictions for PDAC sample TIPC418 from Meng et al.[22] who tested eight TCRs and found none to react to the TIPC418-derived tumor cell line (blue dots, dot size scaled to number of TIL TCR clonotypes). PredicTCR analysis predicted seven of these eight TCRs to be nonreactive (reactivity scores below the Fisher–Jenk natural break threshold, dashed line in plot). Seven additional TCR clonotypes (red dots) predicted to be tumor reactive were cloned for prospective validation of predicTCR. **i**,**j**, Flow cytometry analysis of T cells expressing predicted TIPC418-reactive TCRs cocultured with TIPC418 cells (top) or irrelevant MeWo control cells (bottom) confirmed all seven TCRs to be reactive as assessed by CD107a (**i**) and TNFα (**j**). **k**, The relative frequency and absolute number of recovered TILs for the TCR clonotypes tested in **h**–**j**.

https://doi.org/10.1038/s41587-024-02161-y

reactive score clearly separating TCR clonotypes into a bimodal distribution corresponding to reactive and nonreactive TCRs (Fig. 3a). We tested these predictions by cloning and experimentally validating an additional 29 α/ß TCR chain pairs (representing 22 clonotypes; Fig. 3b), finding that predicTCR$_{50}$ accurately predicted tumor reactivity for 20 out of 22 TCRs (AUC of 0.92 and accuracy of 0.91; Fig. 3c



https://doi.org/10.1038/s41587-024-02161-y

Table 1 | Performance of tumor-reactivity prediction methods on BT21

Method	Threshold	TP	FP	TN	FN	Accuracy	G-mean	AUC
NeoTCR8	0.39	2	0	9	11	0.50	0.15	0.87
Hanada	0.43	13	5	4	0	0.82	0.67	0.72
Caushi	0.46	13	5	4	0	0.77	0.67	0.72
Meng TR30	0.37	12	4	5	1	0.77	0.72	0.85
PredicTCR	0.52	12	1	8	1	0.91	0.91	0.92

The performance of each prediction method on the BT21 test set of 22 TCRs was quantified using accuracy, G-mean and AUC. TP, true positive; FP, false positive; TN, true negative; FN, false negative; and threshold, Fisher–Jenk threshold for calling a TCR as tumor reactive.

Table 2 | PredicTCR does not falsely detect tumor-reactive T cells in PBMC samples from patients with COVID-19 before and after infection

Severity	No. patients	No. cells	PredicTCR		NeoTCR8	
			PFP	CFP	PFP	CFP
Asymptomatic	5	26,093	0	0	3	260
Mild	6	31,911	0	0	2	95
Moderate	2	19,010	0	0	2	380
Severe	12	74,308	0	0	9	670
Mild – Post	2	11,095	0	0	1	98
Moderate – Post	1	6,865	0	0	0	0
Severe – Post	17	70,087	0	0	4	406

Virus-reactive T cells express a transcriptional program that partially overlaps with that of tumor-reactive T cells. PredicTCR does not call any T cells as being tumor reactive in scRNA + VDJ data from PBMCs drawn from patients with COVID-19 (ref. 28), exhibiting better discrimination than gene signature approaches such as the NeoTCR8 signature[23]. Clonotype level analysis not possible as scVDJ-seq was not performed. PFP, the number of patients with false positive cells called as tumor reactive; CFP, the number of cells called as being tumor reactive.

and Table 1). Our CD107a$^+$ threshold for tumor reactivity captured T cells with a broad range of activated phenotypes, with CD8$^+$ BT21-reactive TCRs killing significantly more BT21 cells in bulk culture xCELLigence assays (Extended Data Fig. 5). We were able to recapitulate these results at single-cell resolution by tracking hundreds of individual transgenic effector T cells using miniaturized microwell coculture assays using the Cellply VivaCyte platform (Extended Data Fig. 6). We then implemented the previously published gene signature-based approaches that generate per-cell TCR tumor-reactivity predictions[19,20,23] and used the same clonotype thresholding procedure as for predicTCR$_{50}$ to distinguish tumor-reactive from nonreactive TCR clonotypes. We found that predicTCR$_{50}$ performed considerably better than the signature-based approaches at predicting reactivity in our 22 TCR set: NeoTCR8 (AUC of 0.87 and accuracy of 0.50), Hanada and Caushi (both AUC of 0.77 and accuracy of 0.72) and Meng TR30 (AUC of 0.85 and accuracy of 0.77) as presented in Table 1 (detailed per-clonotype predictions, Uniform Manifold Approximation and Projection (UMAP) plots and ROC curves shown in Extended Data Fig. 7b–e). This was not unexpected given that the signature approaches were derived from other tumor types, while predicTCR$_{50}$ was trained on BT21 data.

Benchmarking predicTCR$_{50}$ false positive rate

Given that 34/50 of the TCRs in the training set and 13/22 of the TCRs in our validation set were tumor reactive, we questioned whether our predicTCR$_{50}$ classifier might have a bias toward calling TCRs as tumor reactive. Published TCR reactivity datasets (such as those used to derive the signature-based approach) typically present more data for tumor-reactive than nonreactive TCRs; this imbalance means that a classifier that calls many TCRs to be tumor reactive would have an apparently high performance. We therefore evaluated the false positive rate of our classifier by analyzing scRNA data from PBMCs of patients with coronavirus disease 2019 (COVID-19). Severe COVID-19 disease is associated with an enrichment of proliferating and effector memory T$_{em}$ populations[28], and previous studies have shown that virus-reactive T cells have a transcriptional signature similar to—but distinct from—tumor-reactive T cells. We found that predicTCR$_{50}$ did not classify any T cells from patients with COVID-19 as tumor reactive (Table 2 and Extended Data Fig. 8), suggesting that predicTCR$_{50}$ is highly specific to tumor-reactive T cells and has a low false positive rate. In contrast, gene signature-based approaches such as NeoTCR8 typically called 1–2% of PBMCs as tumor reactive in a majority of patients with COVID-19, even those recovered from infections with mild symptoms where fewer T cells would be expected to express the virus-reactive signature.

PredicTCR performance generalizes to diverse tumor types

Having shown that our training method did not generate a classifier with a high false positive rate, we created the final version of predicTCR by retraining on all 72 BT21 derived TCRs (1,679 cells) and healthy donor data (111,499 cells). We set out to compare the performance of predicTCR with that of signature approaches using only externally generated data. Given the aforementioned imbalance in validation data, we primarily used the geometric mean (G-mean) of sensitivity (true positive rate) and specificity (true negative rate) to benchmark model performance. We first applied predicTCR to nine PDAC tumors from which tumor cell lines, TIL scRNA + VDJ data and TCR reactivity testing for 118 clonotypes were available[22]. Despite not being trained on PDAC data, we found that predicTCR could accurately predict experimentally determined tumor reactivity as shown in Fig. 3d and detailed in Table 3 (accuracy of 0.88, G-mean of 0.88 and AUC of 0.88). This suggested that predicTCR detects core transcriptional features of tumor-mediated T cell activation that are independent of tumor type. These scores are notably higher than those achieved when applying gene signature approaches including NeoTCR8 (accuracy of 0.47, G-mean of 0.03 and AUC of 0.65; Supplementary Table 2), Hanada (accuracy of 0.77, G-mean of 0.76 and AUC of 0.76; Supplementary Table 3), Caushi (accuracy of 0.54, G-mean of 0.13 and AUC of 0.51; Supplementary Table 4) and the TR30 signature from Meng et al. (accuracy of 0.81, G-mean of 0.81 and AUC of 0.88; Supplementary Table 5), highlighting the increased predictive value of the machine learning-derived classifier.

We next extended our analysis to include additional publicly available TIL scRNA + VDJ datasets from additional tumor types. From Lowery et al., we analyzed a single colorectal metastatic cancer patient (SR4323) for whom both reactive and nonreactive TCRs were available, showing predicTCR accuracy with a G-mean of 0.76 (accuracy of 0.83 and AUC of 0.96; Fig. 3e and Table 3). By comparison, NeoTCR8 performed perfectly on its training dataset with a G-mean of 1.00 (accuracy of 1.00 and AUC of 1.00), whereas the Hanada et al., Caushi et al. and Meng et al. TR30 gene signature-based approaches performed with respective G-means of 0.53 (accuracy of 0.72 and AUC of 0.64), 0.00 (accuracy of 0.61 and AUC of 0.50) and 0.53 (accuracy of 0.72 and AUC of 0.84), suggesting that gene signature-based approaches fail to generalize beyond the tumor type in which they were derived (Supplementary Tables 2–5).

We analyzed three non-small cell lung cancer (NSCLC) samples from Caushi et al. for which TCRs were cloned and tested. Since only ten TCRs were directly tested in these samples, we also included TCRs shown to be neoepitope-reactive based on mutation-associated neoantigen functional expansion or virus-reactive based on viral antigen functional expansion. PredicTCR once again performed well, with a G-mean of 0.87 (accuracy of 0.87 and AUC of 0.94; Fig. 3f and Table 3), better than the Caushi et al. signature derived from these samples, with

Table 3 | Summary of predicTCR TCR tumor-reactivity predictions in diverse cancer types

Patients	Source	Type	Tech.	Validation	Threshold	TP	FP	TN	FN	Accuracy	G-mean	AUC
TIPC249	Meng et al.	PDAC	10X	Cell line	0.39	5	0	0	1	0.83	NA	NA
TIPC262	Meng et al.	PDAC	10X	Cell line	0.39	9	2	6	0	0.88	0.87	0.89
TIPC282	Meng et al.	PDAC	10X	Cell line	0.39	8	1	4	0	0.92	0.89	0.98
TIPC301	Meng et al.	PDAC	10X	Cell line	0.39	0	1	11	3	0.73	0.00	0.61
TIPC309	Meng et al.	PDAC	10X	Cell line	0.39	21	0	13	0	1.00	1.00	1.00
TIPC413	Meng et al.	PDAC	10X	PDX	0.39	1	0	3	1	0.80	0.71	0.67
TIPC416	Meng et al.	PDAC	10X	Cell line	0.39	3	1	2	2	0.63	0.63	0.53
TIPC418	Meng et al.	PDAC	10X	Cell line	0.39	0	1	7	0	0.88	NA	NA
TIPC432	Meng et al.	PDAC	10X	Cell line	0.39	8	0	3	1	0.92	0.94	0.89
TIPC overall	Meng et al.	PDAC	10X	Cell line	NA	55	6	49	8	0.88	0.88	0.88
SR4323	Lowery et al.	Colon-met	10X	TMG	0.45	11	3	4	0	0.83	0.76	0.96
MD01-004	Caushi et al.	NSCLC	10X	Peptide	0.57	8	2	4	0	0.86	0.82	1.00
MD01-005	Caushi et al.	NSCLC	10X	Peptide	0.45	1	0	12	2	0.87	0.58	0.89
MD043-011	Caushi et al.	NSCLC	10X	Peptide	0.54	2	0	0	0	1.00	NA	NA
MD overall	Caushi et al.	NSCLC	10X	Peptide	NA	11	2	16	2	0.87	0.87	0.94
CRI3061	Zheng et al.	GI (PDAC)	SS2*	TMG+Pep	0.16	2	0	4	0	1.00	1.00	1.00
CRI3244	Zheng et al.	GI (PDAC)	SS2*	TMG+Pep	0.24	0	1	7	0	0.88	NA	NA
CRI3281	Zheng et al.	GI (bile duct)	SS2	TMG+Pep	0.15	0	2	2	0	0.50	NA	NA
CRI3395	Zheng et al.	GI (bile duct)	SS2*	TMG+Pep	0.15	1	1	5	0	0.86	0.91	1.00
CRI3571	Zheng et al.	GI (bile duct)	SS2*	TMG+Pep	0.15	1	14	5	0	0.30	0.51	0.74
CRI overall	Zheng et al.	GI	SS2	TMG+Pep	NA	4	18	23	0	0.60	0.78	0.74

The true positive (TP), false positive (FP), true negative (TN) and false negative (FN) clonotype level tumor-reactivity predictions, as well as the resulting accuracy, G-mean and AUC scores for 18 different tumor samples from four studies. The AUC and G-mean calculations were only possible if both reactive and nonreactive TCRs were available, not available (NA) otherwise. Type, Colon-Met, colorectal metastatic cancer; GI, gastrointestinal cancer. Tech., the single-cell sequencing technology used; 10X, 10X Genomics Single Cell Immune Profiling; SS2, Smart-seq2; SS2*, modified Smart-seq2 (scM&T-Seq); validation, the methodology used to test TCR reactivity; cell line, T cell coculture with tumor cell line, where PDX is T cell coculture with patient-derived xenograft, TMG is T cell coculture with antigen presenter cells expressing a tandem minigene encoding tumor neoantigens and Pep is T cell coculture with peptide-pulsed antigen presenter cells; and threshold, Fisher–Jenk threshold for calling a TCR as tumor reactive.

which we observed a G-mean of 0.75 (accuracy of 0.74 and AUC of 0.83; Supplementary Table 4). The Hanada et al. signature was also derived from NSCLC samples, and as expected, it performed similarly to the Caushi et al. signature with a G-mean of 0.76 (accuracy of 0.81 and AUC of 0.86; Supplementary Table 3). NeoTCR8, on the other hand, was not predictive with a G-mean of 0.00 (accuracy of 0.58 and AUC of 0.50; Supplementary Table 2), while surprisingly the TR30 signature derived from PDAC samples performed well with a G-mean of 0.79 (accuracy of 0.77 and AUC of 0.98).

Finally, we analyzed five gastrointestinal cancer samples generated by Zheng et al.[21] using Smart-seq2 (ref. 29), which contain an order of magnitude fewer cells (mean 328 cells per sample) than other external datasets generated using the 10x Genomics platform[19,20,23]. Notably, this dataset contained testing data for both CD4 and CD8 T cells; however, as published signature-based prediction approaches focused on CD8 cells, we restricted our comparisons to only the CD8 T cell data. For three samples, predicTCR performed with high accuracy (Fig. 3g and Table 3), while for two samples accuracy was reduced, leading to an overall G-mean of 0.78 (accuracy of 0.60 and AUC of 0.74). All other gene signature approaches were not predictive, with G-means of 0.00 (accuracy of 0.91 and AUC of 0.50; Supplementary Table 2), 0.41 (accuracy of 0.24 and AUC of 0.54; Supplementary Table 3), 0.16 (accuracy of 0.11 and AUC of 0.57; Supplementary Table 4) and 0.41 (accuracy of 0.24 and AUC of 0.52; Supplementary Table 5) for NeoTCR8, Hanada, Caushi and Meng TR30 gene signatures, respectively. These results suggest that predicTCR is applicable to datasets with few cells (probably to include tumor biopsies that can be more easily obtained than resection material), and sequenced at lower cost due to the reduced cell number.

Having demonstrated the generalizability of predicTCR, we set out to experimentally validate a number of TCRs predicted to be tumor reactive in a different tumor type. Meng et al. processed a PDAC sample (TIPC418) and tested 12 TCR clonotypes, eight of which were found to be nonreactive and four of which showed weak reactivity[22]. PredicTCR analysis identified many other TCR clonotypes with a high chance of being tumor reactive (Fig. 3h). From these, we selected seven new TCR clonotypes expressed in multiple TILs (Source data) and confirmed that all seven showed reactivity against the TIPC418 PDAC line as assessed by flow cytometry-based quantification of CD107a and TNFα but not against a negative control MeWo cell line (Fig. 3i,j). In this sample, many of the nonreactive TCRs were present at higher frequencies in the TIL population than the reactive clonotypes (Fig. 3k); we found that the two most frequent clonotypes shared a CDR3 α sequence that has been reported to bind to a cytomegalovirus-derived epitope in VDJdb[30], confirming the utility of predicTCR in identifying TCRs for personalized cell therapies.

Discussion

Here, we present predicTCR, the first automated classifier of tumor TCR reactivity capable of highly accurate identification of tumor-reactive TCR clonotypes in TILs derived from diverse cancer types through the use of machine learning models combined with deterministic thresholding. We show that through careful sample choice, generation of a large, high-confidence TCR reactivity dataset and inclusion of extensive negative training data, an accurate classifier can be generated. In contrast to previous approaches using differential gene expression to elucidate a gene signature specific to one tumor type, predicTCR enables rapid, antigen-agnostic identification of tumor-reactive TCRs

in diverse tumor types—the first step in manufacturing personalized TCR-transgenic T cell cancer therapies.

The majority of signature-based approaches rely on gene set enrichment analysis of clustered tumor-reactive T cells, and thus identify genes that are upregulated in the *CXCL13* expressing cluster that we have previously shown to contain tumor-reactive infiltrating T cells[15]. However, *CXCL13* expression does not always define a discrete population of T cells (see the example in Extended Data Fig. 9), as cell clustering is highly dependent on upstream processing methods such as normalization[31], the type of clustering algorithm used and the number of cells in a particular dataset[32]. Furthermore, tumor-reactive TCRs may be expressed in T cells of diverse phenotypes (including memory and exhausted populations): here, cluster-based approaches struggle to interpret genes that are expressed across clusters but which have context-specific predictive value that can be discriminated by machine learning. Finally, clustering approaches require manual verification and annotation to achieve optimal results, making automation difficult. This particularly affects small datasets such as those generated with Smart-seq2, for which cluster-based gene signatures did not detect reactive T cell clusters in as many as six out of ten patients (Zheng et al.[21]).

We interrogated the predicTCR classifier using explainable AI SHAP to determine the key genes marking tumor reactivity in T cells. While the known reactivity marker *CXCL13* contributed the most to our classifier for prediction (Extended Data Fig. 10), of the two next best genes, AC243829.4 was only identified by Caushi et al., while LINC02099 was completely absent from signature approaches. AC243829.4 has recently been reported to correlate with the presence of immune cells in the tumor microenvironment in clear cell renal cell carcinoma, and is associated with positive patient prognosis[33], possibly by regulating the expression of the inflammatory cytokine *CCL3* (ref. 34). Notably, the relationship between expression of LINC02099 and tumor reactivity is not linear, which we believe is the result of LINC02099 being identified as the hub of a large long noncoding RNA–messenger RNA (mRNA) regulatory network in a breast cancer study[35], giving rise to complex interaction effects that can be best determined by machine learning approaches.

Given the high cost of manufacturing personalized TCR-mediated cell therapies under GMP or GMP-like conditions, only a limited number of TCRs can be manufactured per patient, so it will be important to avoid manufacturing nontumor-reactive TCRs (that is, false positive predictions). As predicTCR generates per-cell reactivity predictions, predicted reactive TCR clonotypes can be ranked by their mean reactivity score: for BT21 picking TCR clonotypes with reactivity scores above the 95th percentile threshold would exclude the one false positive predicTCR prediction we obtained by using a binary reactivity threshold (Extended Data Fig. 7). Among the TCR clonotypes exceeding a given threshold, the most frequent TCR clonotypes can be prioritized as having a more reliable reactivity score, as well as showing evidence of antigen-driven expansion. Finally, complementary analysis of the CDR3 repertoire can assist in picking between TCRs of similar frequency and score, such as identifying clusters of similar CDR3 sequences that are statistically unlikely to occur in naive repertoires using tools such as ALICE (antigen-specific lymphocyte identification by clustering of expanded sequences)[36]. We illustrate a cluster of TCRs in the BT21 TIL repertoire that have convergently recombined the tumor-reactive CDR3 ß sequence 'CASSLGGASYEQYF' in Supplementary Table 6. Of less importance in a translational context is the single false negative reactivity prediction made by predicTCR in the BT21 test set. We speculate that this might reflect a bona fide tumor-reactive TCR which could not be validated using the BT21 cell line, either due to transcriptional changes occurring to the cell line during adaptation to cell culture conditions resulting in downregulation of the TCR's target antigen or due to the BT21 cell line lacking SNVs found in the original tumor. We note that in general, the performance of predicTCR is better on TCR tumor reactivity datasets generated using a tumor cell line as the T cell target. Tumor cell lines recapitulate the diversity of potential TCR target antigens found in the original tumor, including tumor-associated antigens, posttranslationally modified antigens and neoepitopes derived from cryptic splicing or the dark proteome, some of which are hard to capture using tandem minigene (TMG) assays. It is therefore possible that the higher false positive prediction rate exhibited by predicTCR when analyzing external samples generated using TMG assays to validate TCR reactivity actually reflects a higher false negative TCR reactivity testing rate in the source assays.

In general, predicTCR predicted more tumor-reactive TCR clonotypes for each sample than could be practicably manufactured for a personalized cell therapy. We found this to be the case even for PDAC samples, which are generally considered to be 'cold tumors' with a low tumor mutation burden and limited T cell infiltration, which may therefore be refractory to conventional TIL therapies[2]. PredicTCR predictions can be refined with accessory analyses, such as computational prediction of TCR avidity that has been shown to enrich for neoantigen-specific TCRs over tumor-associated antigen-specific TCRs[37]. Optimally, a minimal panel of computationally predicted tumor-reactive TCRs will advance to experimental resolution of the target epitope using sensitive, high dynamic range reporters of T cell activation[38]. Such analyses might be further informed by computational reconstruction of tumor heterogeneity to identify clonal or near-clonal tumor mutations[39]; TCRs reactive to these targets are most likely to result in tumor clearance. Combining these new computational and experimental tools will allow for the creation of a validated patient-derived cell therapy product targeting diverse, tumor-specific, clonal antigens at lower cost than current screening. However, for aggressive cancers in which patient survival is short, the rapid sample-to-vein turnaround enabled by predicTCR would allow for the creation of a personalized cell therapy product in an entirely antigen-agnostic fashion. Although the TCRs contained in such a product would target unknown antigens, given that autologous TCRs have undergone thymic selection, they pose little risk to patients, and targeting subclonal (that is nonoptimal) tumor antigens may yet offer patients clinical benefit by nucleating an immune cascade and epitope spreading effects. Furthermore, resolving the target epitope of a TCR from TCR sequence alone is a rapidly advancing field[40] and we believe that by pairing tumor mutanome data with predicTCR reactivity predictions, datasets with dramatically reduced numbers of possible TCR–epitope interactions can be generated, serving to train and test TCR–epitope prediction tools, which will themselves allow for the training of TCR reactivity classifiers with ever higher predictive accuracy. As the costs of TCR synthesis fall and validated scRNA + VDJ-seq datasets become more widely available, it will become possible to generate increasingly large training datasets, ensuring that future classifiers can identify tumor-reactive TCRs (or specialized subsets thereof) with even greater accuracy.

In conclusion, we believe that accurate machine learning classifiers such as predicTCR will accelerate the realization of personalized T cell-mediated transgenic cell therapies by reducing overall sample-to-vein turnaround times and increasing the likelihood of therapy delivery before tumor progression, while reducing the costs that currently limit implementation.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41587-024-02161-y.

References

1. Rohaan, M. W. et al. Tumor-infiltrating lymphocyte therapy or ipilimumab in advanced melanoma. *N. Engl. J. Med.* **387**, 2113–2125 (2022).

2. Monberg, T. J., Borch, T. H., Svane, I. M. & Donia, M. TIL therapy: facts and hopes. *Clin. Cancer Res.* **29**, 3275–3283 (2023).
3. Cohen, C. J. et al. Isolation of neoantigen-specific T cells from tumor and peripheral lymphocytes. *J. Clin. Invest.* **125**, 3981–3991 (2015).
4. Crompton, J. G., Sukumar, M. & Restifo, N. P. Uncoupling T cell expansion from effector differentiation in cell-based immunotherapy. *Immunol. Rev.* **257**, 264–276 (2014).
5. Poschke, I. C. et al. The outcome of ex vivo TIL expansion is highly influenced by spatial heterogeneity of the tumor T cell repertoire and differences in intrinsic in vitro growth capacity between T cell clones. *Clin. Cancer Res.* **26**, 4289–4301 (2020).
6. Zacharakis, N. et al. Immune recognition of somatic mutations leading to complete durable regression in metastatic breast cancer. *Nat. Med.* **24**, 724–730 (2018).
7. Wang, B. et al. Generation of hypoimmunogenic T cells from genetically engineered allogeneic human induced pluripotent stem cells. *Nat. Biomed. Eng.* **5**, 429–440 (2021).
8. Simoni, Y. et al. Bystander CD8$^+$ T cells are abundant and phenotypically distinct in human tumour infiltrates. *Nature* **557**, 575–579 (2018).
9. Hundal, J. et al. PVACtools: a computational toolkit to identify and visualize cancer neoantigens. *Cancer Immunol. Res.* **8**, 409–420 (2020).
10. Shah, N. M. et al. Pan-cancer analysis identifies tumor-specific antigens derived from transposable elements. *Nat. Genet.* **55**, 631–639 (2023).
11. Bartok, O. et al. Anti-tumour immunity induces aberrant peptide presentation in melanoma. *Nature* **590**, 332–337 (2021).
12. Kacen, A. et al. Post-translational modifications reshape the antigenic landscape of the MHC I immunopeptidome in tumors. *Nat. Biotechnol.* **41**, 239–251 (2023).
13. Ouspenskaia, T. et al. Unannotated proteins expand the MHC-I-restricted immunopeptidome in cancer. *Nat. Biotechnol.* **40**, 209–217 (2022).
14. Naghavian, R. et al. Microbial peptides activate tumour-infiltrating lymphocytes in glioblastoma. *Nature* **617**, 807–817 (2023).
15. Platten, M. et al. A vaccine targeting mutant IDH1 in newly diagnosed glioma. *Nature* **592**, 463–468 (2021).
16. Oliveira, G. et al. Phenotype, specificity and avidity of antitumour CD8$^+$ T cells in melanoma. *Nature* **596**, 119–125 (2021).
17. Oliveira, G. et al. Landscape of helper and regulatory antitumour CD4$^+$ T cells in melanoma. *Nature* **605**, 532–538 (2022).
18. Veatch, J. R. et al. Neoantigen-specific CD4$^+$ T cells in human melanoma have diverse differentiation states and correlate with CD8$^+$ T cell, macrophage, and B cell function. *Cancer Cell* **40**, 393–409.e9 (2022).
19. Caushi, J. X. et al. Transcriptional programs of neoantigen-specific TIL in anti-PD-1-treated lung cancers. *Nature* **596**, 126–132 (2021).
20. Hanada, K. et al. A phenotypic signature that identifies neoantigen-reactive T cells in fresh human lung cancers. *Cancer Cell* **40**, 479–493.e6 (2022).
21. Zheng, C. et al. Transcriptomic profiles of neoantigen-reactive T cells in human gastrointestinal cancers. *Cancer Cell* **40**, 410–423.e7 (2022).
22. Meng, Z. et al. Transcriptome-based identification of tumor-reactive and bystander CD8$^+$ T cell receptor clonotypes in human pancreatic cancer. *Sci. Transl. Med.* **15**, 1–19 (2023).
23. Lowery, F. J. et al. Molecular signatures of antitumor neoantigen-reactive T cells from metastatic human cancers. *Science* **375**, 877–884 (2022).
24. Joshi, K. et al. Spatial heterogeneity of the T cell receptor repertoire reflects the mutational landscape in lung cancer. *Nat. Med.* **25**, 1549–1559 (2019).
25. Hu, Z., Li, Z., Ma, Z. & Curtis, C. Multi-cancer analysis of clonality and the timing of systemic spread in paired primary tumors and metastases. *Nat. Genet.* **52**, 701–708 (2020).
26. Chen, T. & Guestrin, C. XGBoost: a scalable tree boosting system. *Proc. 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* 785–794 (ACM, 2016).
27. Lundberg, S. M. & Lee, S.-I. in *Advances in Neural Information Processing Systems* Vol. 30 (eds Guyon, I. et al.) (Curran Associates, 2017).
28. Yoshida, M. et al. Local and systemic responses to SARS-CoV-2 infection in children and adults. *Nature* **602**, 321–327 (2022).
29. Picelli, S. et al. Smart-seq2 for sensitive full-length transcriptome profiling in single cells. *Nat. Methods* **10**, 1096–1100 (2013).
30. Shugay, M. et al. VDJdb: a curated database of T cell receptor sequences with known antigen specificity. *Nucleic Acids Res.* **129**, 170–177 (2017).
31. Ahlmann-Eltze, C. & Huber, W. Comparison of transformations for single-cell RNA-seq data. *Nat. Methods* **20**, 665–672 (2023).
32. Yu, L., Cao, Y., Yang, J. Y. H. & Yang, P. Benchmarking clustering algorithms on estimating the number of cell types from single-cell RNA-sequencing data. *Genome Biol.* **23**, 49 (2022).
33. Wang, Z. et al. Identification and verification of immune subtype-related lncRNAs in clear cell renal cell carcinoma. *Front. Oncol.* **12**, 369–373 (2022).
34. Song, G. Y. et al. Differential expression profiles and functional analysis of long non-coding RNAs in calcific aortic valve disease. *BMC Cardiovasc. Disord.* **23**, 1–13 (2023).
35. Wu, W. et al. Tissue-specific co-expression of long non-coding and coding RNAs associated with breast cancer. *Sci. Rep.* **6**, 1–13 (2016).
36. Pogorelyy, M. V. et al. Detecting T cell receptors involved in immune responses from single repertoire snapshots. *PLoS Biol.* **17**, 1–13 (2019).
37. Schmidt, J. et al. Neoantigen-specific CD8 T cells with high structural avidity preferentially reside in and eliminate tumors. *Nat. Commun.* **14**, 3188 (2023).
38. Schmid, T. et al. T-FINDER: a highly sensitive, pan-HLA platform for functional T cell receptor and ligand discovery. *Sci. Adv.* **10**, adk3060 (2024).
39. Grigoriadis, K. et al. CONIPHER: a computational framework for scalable phylogenetic reconstruction with error correction. *Protoc. Exch.* **562**, 833–845 (2023).
40. Jokinen, E. et al. TCRconv: predicting recognition between T cell receptors and epitopes using contextualized motifs. *Bioinformatics* **39**, btac788 (2023).

Methods

Sample and patient consent

Patient BT21, a 62-year-old male previously diagnosed with melanoma, was treated for a brain metastasis at the University Hospital Mannheim following written consent. The patient was not financially compensated for participation. The study was approved by the institutional review board (Ethikkommission 2019-643N).

Processing of tumor samples for sequencing

Freshly resected brain tumor tissue was obtained from the University Hospital in Mannheim. The patient gave informed written consent before sample collection. Tissue was transported on ice in phosphate-buffered saline (PBS) (Sigma-Aldrich) and processed within 3 h of resection by dissection into into small pieces (2 × 2 × 2 mm). Individual tumor pieces were snap frozen and stored at −80 °C before extracting DNA and RNA for sequencing. The whole-exome library was prepared using SureSelect Human All Exon V7 (5191-4028, Agilent) and the RNA sequencing library was prepared using Ultra Low Input RNA-Seq from TakaraBio. Both were sequenced using NovaSeq 6000 (2×100 bp). DNA isolated from PBMCs from patient BT21 was included as the whole-exome reference sample.

The remaining tumor pieces were gently mashed through a 100 µm cell strainer using the back side of a syringe plunger to generate a single-cell suspension. To generate a tumor cell line, a portion of the single-cell suspension was spun down (350*g*, 5 min, room temperature) and resuspended in Dulbecco's modified Eagle medium/F12 (Gibco) supplemented with 1× penicillin–streptomycin (Sigma), 1× B27 supplement (Thermo Fisher), 20 ng ml^{-1} epidermal growth factor (236-EG, R&D Systems) and 20 ng ml^{-1} fibroblast growth factor (13256-029, Thermo Fisher). Cells were placed in a 37 °C CO_2 incubator where they started to grow as spheroids. Cells were subsequently transferred into Roswell Park Memorial Institute (RPMI)-1640 media (Sigma) supplemented with penicillin–streptomycin and 10% fetal bovine serum (FBS), whereupon they grew as a monolayer. Cells were split with accutase (A1110501, Thermo Fisher) when appropriate during establishment of the robustly growing tumor cell line.

The remaining single-cell suspension was filtered through a 70 µm cell strainer, myelin was removed using myelin removal beads II (130-096-433, Miltenyi) and LS columns (130-042-401, Miltenyi) according to the manufacturer's protocol, and aliquots of the single-cell suspension were cryopreserved as described for PBMCs. Thawed aliquots were used for fluorescence-activated cell sorting (FACS)-based enrichment of T cells (CD3$^+$ and CD45$^+$) and prepared for sequencing using Chromium Single Cell V(D)J Reagent kit v1.1 chemistry (PN-1000006, PN-1000020, PN-1000005 and PN-120262, 10X Genomics) according to the manufacturer's protocol. The constructed scVDJ library and scGEX libraries were sequenced using the NovaSeq 6000 platform (Illumina).

Exome sequence variant calling

Variant calling was performed by the German Cancer Research Center Omics Data Core Facility using previously described pipelines[41]. Briefly: exome sequencing was performed on DNA extracted from PBMCs, tumor and the tumor cell line. SNVs were called relative to the human genome reference sequence GRCh37, and tumor and cell line SNVs determined by subtracting germline SNVs present in the PBMC sample using the One Touch Pipeline[42].

In silico HLA typing from bulk RNA-seq data

For in silico human leukocyte antigen (HLA) typing on paired fastq files from bulk RNA-seq analysis, arcasHLA[43] was used to perform in silico HLA typing on paired fastq files from bulk RNA-seq analysis.

Recovery of TCR sequences from bulk RNA-seq data

We used TRUST4 to reconstruct unpaired α and ß TCR chain sequences from within the bulk RNA-seq data as described by Song et al.[44].

Generation of TCR in vitro-transcribed mRNA constructs

Cell Ranger-derived TCR clonotype data were processed in R using tidyverse functions[45]. VDJ regions of TCRs were ordered as synthetic DNA fragments from Twist Biosciences and cloned in 96-well format as chimeric TCRs, using murine TRAC or TRBC constant region sequences that had been further modified to include an additional disulfide bond to improve stability and avoid mismatches with the endogenous human TCR after transduction into human T cells[46,47]. As negative controls, we cloned two TCRs targeting HLA-A*02:01 restricted epitopes of MART1 (DMF5 TCR: CDR3α CAVNFGGGKLIF and CDR3β CASSLSFGTEAFF) or influenza (CDR3α CAVSESPFGNEKLTF and CDR3β CASSSTGLPYGYTF). For in vitro transcription, RNA-mediated expression TCR constructs were PCR amplified using a primer to add a T7 promoter, and the resulting PCR product used as a template for the T7 mScript Standard mRNA Production System (CELLSCRIPT C-MSC11610). mRNA was m^{7}G capped and enzymatically polyadenylated following the manufacturer's instructions. For TCR killing assays, TCR constructs were subcloned into S/MAR nanovectors using classical molecular biology techniques as previously described[48].

Isolation and expansion of healthy donor PBMCs

PBMCs from healthy donors were isolated from heparinized blood. In short, 15.5 ml of Ficoll Paque Plus Media (Cytiva) was loaded per Leucosep tube (Greiner Bio-One) and spun down. After adding 3 ml of PBS (Sigma), up to 25 ml of blood was loaded on top and a density-gradient centrifugation was performed at 800*g* (acceleration 4 and deceleration 3). After collection of the interphase, PBMCs were washed twice with PBS and frozen in a controlled rate freezing device at −80 °C in 50% freezing medium A (60% X-Vivo 20 and 40% fetal calf serum) and 50% medium B (80% fetal calf serum and 20% dimethylsulfoxide). Cells were stored in liquid nitrogen at −140 °C until further analysis.

The rapid expansion protocol was used to expand T cells. PBMCs from three independent donors were irradiated at 40 Gy using a Gammacell 1000 (AECL) irradiation device to serve as feeder cells. Then, 1×10^7 cells from each donor were pooled together, cells were spun down (400*g*, 10 min, room temperature) and resuspended in rapid expansion protocol media (X-Vivo15 (Lonza, BE02-060Q), 2% human AB serum (H4522-100ML, Sigma-Aldrich), 2.5 µg ml^{-1} Fungizone (15290-018, Gibco), 20 µg ml^{-1} gentamicin (2475.1, Roth), 100 IU ml^{-1} penicillin and 100 µg ml^{-1} streptomycin (15140122, Life Technologies)). Next, 150,000 PBMCs were plated into a standing T25 flask and 666 ng of OKT-3 antibody (Life Technologies, 16-0037-85) was added to the culture and the flask was topped up to a total volume of 20 ml. The next day, 5 ml of X-Vivo15 supplemented with 2% AB serum containing 7,500 IU interleukin-2 (IL-2) was added to the culture. Three days later, 12.5 ml of medium was removed and replaced with 12.5 ml of X-Vivo15 supplemented with 2% AB serum containing 600 IU ml^{-1} IL-2.

Melan A expression

Melan A expression was confirmed using anti-Melan A-FITC (cat. no. sc-20032, clone A103, Santa Cruz Technology), diluted at 1:10.

TCR reactivity screening via flow cytometry

TCR-encoding RNA was electroporated into expanded healthy donor PBMCs using the Lonza 4D-Nucleofector (program EO-115, solution P3 supplemented according to the manufacturer's recommendations), which were plated into 48-well plates containing TexMACS media (130-097-196, Miltenyi) supplemented with 2% human AB serum. At 18–24 h after electroporation, cells were collected and 50 IU ml^{-1} benzonase (YCP1200-50KU, Speed BioSystems) was added to avoid cell clumping. TCR expression levels were measured via flow cytometry with markers including fixable viability dye (AF700, BD), CD3 (clone HIT3A, BV510, BD) and mTCRb (clone H57-597, PE, Biolegend).

To assess TCR reactivity, a total of 150,000 T cells and 75,000 cells of the patient-autologous tumor cell line were cocultured in U-bottom

96-well plates in a total volume of 200 µ. Wells with only T cells, or T cells and TransAct beads (130-111-160, Miltenyi) were used as negative and positive controls, respectively. Then, 5 µl of CD107a FACS antibody (REF 561343, BD) was added per well. After 1 h of coculture, GolgiPlug and GolgiStop (555029 and 554724, BD) were added to reach a 1:1,000 dilution, and after four additional hours of coculture cells were used for flow cytometry analysis. Markers included fixable viability dye (AF700, 1:1,000 dilution, eBioscience), CD3 (clone HIT3A, BV510, 1:20 dilution, BD), mTCRb (clone H57-597, 1:50 dilution, PE, Biolegend) and TNFa (clone MAb11, BV711, 1:10 dilution, Biolegend). Samples were acquired on a FACS Lyric device and flow cytometry data were analyzed using FlowJo software, v10.6.2 (FlowJo LLC).

TCRs were classified as reactive or nonreactive based on flow cytometry data acquired after coculture. The percentage of CD107a positive cells (%CD107a) was quantified by gating on viable $CD3^+$ $mTCR\beta^+$ singlets. TCRs were included in the analysis if the mTCRβ expression was >2%.

The %CD107a signal per TCR after coculture with the cell line ('TCR versus cell line') or after running the coculture assay without stimulation ('TCR, unstimulated') was corrected for background by calculating

$$(\%\mathrm{CD107a}_{\mathrm{TCRvscellline}} - \%\mathrm{CD107a}_{\mathrm{TCR,unstimulated}}) - (\%\mathrm{CD107a}_{\mathrm{Mockvscellline}} - \%\mathrm{CD107a}_{\mathrm{Mock,unstimulated}})$$

where mock refers to expanded T cells electroporated without TCR-encoding RNA. TCRs were classified as reactive if the background corrected %CD107a signal per TCR was larger than 2× the standard deviation of the $\%CD107a^+$ signal measured in all samples without stimulation (1× s.d. of 0.34%). Where a TCR clonotype expressed two α chains, data are presented for the α chain resulting in the %CD107a expression (that is, the functional pair).

TCR reactivity screening via xCELLigence real-time killing assays

Primary human $CD3^+$ cells were isolated from healthy donor volunteers using the Pan T cell isolation kit from Miltenyi Biotec according to the manufacturer's instructions. The isolated T cells were then activated for 3 days using the human T Cell TransAct kit (Miltenyi Biotec) according to the manufacturer's instructions and cultured in TexMACS medium from Miltenyi Biotec supplemented with IL-7 and IL-15, both at a final concentration of 10 ng ml^{-1}, at a concentration of 1×10^6 cells ml^{-1}. 3 days post activation 2×10^6 cells were washed and resuspended in 20 µl of primary P3 solution (Lonza), mixed with 2 µg of S/MAR DNA nanovectors and pulsed with the FI-115 pulsing code using the Lonza 4D-Nucleofector.

Primary human T cells were collected, washed two times and resuspended in FACS buffer (PBS containing 1% of FBS). TCR expression was detected by flow cytometry and T cells were stained with a PE-conjugated antibody (clone H57-597, PE, Biolegend) for 30 min on ice in the dark. Dead cells were excluded by 4,6-diamidino-2-phenylindole gating and alive TCR^+ cells were gated. Data analysis was performed using FlowJo software.

A real-time killing assay using the xCELLigence was performed. Briefly, BT21 tumor cells were seeded on a 96-well plate (3×10^4 cells per well) and incubated for 24 h. Transgenic T cells were added at an effector–target cell ratio of 2:1 and co-incubated at 37 °C in RPMI 10% medium for 24 h. Cell growth was then monitored for 24 h.

TCR reactivity screening via cell-mediated cytotoxicity

Analysis of transgenic TCR cell cytotoxicity at microfluidic scale was carried out on the VivaCyte platform (Cellply) loaded with a CC-Array microfluidic device based on a modified version of the open-microwell technology[49]. The CC-Array contains 16 lanes, each lane comprising 1,200 microwells where effector and target cells can interact. Lower microfluidic channels under the microwell array of the CC-Array device were initially preloaded with 6% gelatin methacryloyl hydrogel (900622, Sigma-Aldrich) in PBS and the gel was polymerized with an ultraviolet lamp. BT21 target cells were prestained with CellTracker Blue CMAC Dye (C2110, Thermo Fisher, Invitrogen). Transgenic T cells and BT21 target cells were resuspended in 100% FBS (10270106, Thermo Fisher, Gibco) and loaded on the upper channels of the CC-Array device, resulting in the formation of cocultures on the bottom part of the microwell at the interface between the liquid and the underlying gelatin methacryloyl layer. Each lane was loaded with T cells expressing a single TCR. After cell delivery, a solution of RPMI-1640 (R0883, Sigma-Aldrich) and propidium iodide (P3566, Thermo Fisher) was then delivered into the microchannels and the microfluidic design allowed to rapidly exchange media in the microwells without displacing the cells. The CC-Array device was maintained at 37 °C, 5% CO_2 and >90% relative humidity in the VivaCyte instrument for the duration of the assay and fluorescence images were acquired every 2 h for 12 h.

An automated analysis of the images was carried out by the VivaCyte software featuring a pretrained deep learning method[50] to detect target cell cytoplasm. Nine hundred microwells were imaged per microchannel by acquiring 20 subarrays per microchannel. Cell viability was quantified as the frequency of cells stained with CMAC and not stained with propidium iodide.

ScRNA-seq analysis

Fastq files from sequenced TIL samples were processed using 10X Genomics' Cellranger v6.1.2 (ref. 51) and count matrices are imported into R v4.1 (ref. 52). Briefly, SoupX[53] was used to removed background noise and miQC[54] used to remove poor quality or degraded cells (that can be identified as having an unusually high mitochondrial gene expression). Cells with an 'RNA count' <1,200 and 'Feature count' <500 were excluded from further analysis.

Healthy PBMC datasets

PBMC datasets enriched with T cells from healthy donors were obtained as follows: a single healthy donor PBMC sample from 10X Genomics[55], two donors from Szabo et al.[56] and seven donors from Gao et al.[57]. In total, data from 111,499 T cells were obtained.

PredicTCR classifier training

All scRNA count data from both internally and externally generated datasets were normalized using the 'sctransform' method as implemented in Seurat v4 (ref. 58), resulting in a gene–cell matrix of Pearson residuals that was used as the model input. TCR reactivity was converted to a binary value from the CD107a flow cytometric quantification as described above; all healthy donor PBMCs were assumed to be nonreactive. The model was trained using scRNA + VDJ-seq data from healthy donors (111,499 cells) plus data from experimentally validated BT21 derived TCRs for $predicTCR_{50}$ (1,461 cells) or predicTCR (1,679 cells) as appropriate. Data were imported in Python (v3.9.16) using pandas (v2.0.2) for preprocessing before training with xgboost (v1.7.4). Due to the scRNA data having many dropouts, we performed hyperparameter tuning before feature selection. The XGBoost hyperparameters 'colsample_bytree', 'gamma', 'learning_rate', 'max_delta_step', 'max_depth', 'min_child_weight', 'n_estimators', 'alpha', 'lambda', 'scale_pos_weight' and 'subsample' were tuned by Bayesian optimization using scikit-optimize (v0.9.0)[59] with ten stratified *k*-fold cross-validations to generate an intermediate classifier model. Due to the imbalanced nature of the training dataset, particular attention was put on optimizing data weighting ('scale_pos_weight'). We used 70% of the data as training data, and the remaining 30% as testing dataset for hyperparameter training. To prevent overfitting to the BT21 training data, we simplified the intermediate classifier using SHAP[27] to identify the key genes contributing to the model. The final predicTCR classifier was then trained on the top 100 SHAP features and hyperparameters were again optimized as before.

Prediction of tumor-reactive T cells using predicTCR

External datasets used to validate predicTCR were downloaded and the raw data preprocessed as described above. The prediction probability for each cell was averaged for each clonotype and the subsequent prediction probability for each clonotype was used to calculate the AUC using pROC. The threshold used to classify TCR reactivity was determined using Fisher–Jenk natural break optimization as implemented in jenkspy. The confusion matrix and accuracy of the resulting prediction were then calculated using caret (v6.0-94), and G-mean (the square root of sensitivity and specificity) was calculated using the output of caret.

Prediction of tumor-reactive T cells using the NeoTCR8 gene signature

Predictions using the NeoTCR8 gene signature were performed as described in Lowery et al. Briefly, the raw gene count matrix was imported into R and scGSEA (using GSVA package, v1.46.0) was performed using the signature gene list (NeoTCR8) obtained from Lowery et al.[23]. Cluster(s) that correspond to 0.95 percentile expression were designated as reactive. A reactive score was calculated using the ratio of predicted reactive cell to the total number of cells for each clonotype. The AUC was then calculated based on this probability score using pROC. To make direct comparisons with the performance of predicTCR, we applied the same Fisher–Jenk optimization to determine the threshold for distinguishing between reactive and nonreactive TCR clonotypes on the basis of the reactive score.

Prediction of tumor-reactive T cells using the Hanada et al. gene signature

Signature analysis using the Hanada et al. gene signature was performed as described in Hanada et al. Briefly, the raw gene count matrix was imported into R and the score was calculated by adding the genes that contributed positively to the signature and minus the genes that contributed negatively to the signature. Cells that were positive for the signature were called as (neoantigen) reactive. A reactive score was calculated and a minimum threshold for tumor reactivity was determined using Fisher–Jenk optimization as described above.

Prediction of tumor-reactive T cells using the Caushi et al. gene signature

Signature analysis using the Caushi et al. gene signature was performed similarly to Caushi et al. Briefly, the raw gene count was imported into R and analyzed using Seurat. Seurat was used to normalize the raw count data, then using 'AddModuleScore', a signature score was calculated using the mutation-associated neoantigen functional expansion genes. Cells that were positive for the signature were called as reactive. A reactive score was calculated and a minimum threshold for tumor reactivity was determined using Fisher–Jenk optimization as described above.

Prediction of tumor-reactive T cells using the Meng et al. TR30 gene signature

Signature analysis using Meng TR30 gene signature was performed as described Meng et al.[22]. Briefly, the raw gene count was imported into R and analyzed using Seurat. Seurat was used to normalize the raw count data; then the TR30 signature was computed using the UCell package (v2.2)[60]. The mean of the TR30 signature score was then calculated for each TCR clonotype and termed the Meng TR30 score. The minimum threshold for tumor reactivity was similarly determined using Fisher–Jenk optimization.

Material availability

The use of the primary tumor cell lines specified in this manuscript is restricted by patient informed consent and institutional review board approval to this study.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

Single-cell sequencing data for BT21 have been deposited in National Center for Biotechnology Information BioProject with accession code PRJNA985415. External datasets were obtained from GSE123139, GSE173351 and phs002748.v1.p1. Source data are provided with this paper.

Code availability

The code is hosted on Zenodo (https://zenodo.org/records/8059129, https://doi.org/10.5281/zenodo.8059129) and will be made available to academic researchers on request.

References

41. Campbell, P. J. et al. Pan-cancer analysis of whole genomes. *Nature* **578**, 82–93 (2020).
42. Reisinger, E. et al. OTP: an automatized system for managing and processing NGS data. *J. Biotechnol.* **261**, 53–62 (2017).
43. Orenbuch, R. et al. ArcasHLA: high-resolution HLA typing from RNAseq. *Bioinformatics* **36**, 33–40 (2020).
44. Song, L. et al. TRUST4: immune repertoire reconstruction from bulk and single-cell RNA-seq data. *Nat. Methods* **18**, 627–630 (2021).
45. Wickham, H. et al. Welcome to the Tidyverse. *J. Open Source Softw.* **4**, 1686 (2019).
46. Cohen, C. J., Zhao, Y., Zheng, Z., Rosenberg, S. A. & Morgan, R. A. 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 (2006).
47. Cohen, C. J. et al. Enhanced antitumor activity of T cells engineered to express T cell receptors with a second disulfide bond. *Cancer Res.* **67**, 3898–3903 (2007).
48. Bozza, M. et al. A nonviral, nonintegrating DNA nanovector platform for the safe, rapid, and persistent manufacture of recombinant T cells. *Sci. Adv.* **7**, eabf1333 (2021).
49. Bocchi, M. et al. Inverted open microwells for cell trapping, cell aggregate formation and parallel recovery of live cells. *Lab Chip* **12**, 3168–3176 (2012).
50. Schmidt, U., Weigert, M., Broaddus, C. & Myers, G. *Lecture Notes in Computer Science* Vol. 11071 (Springer, 2018).
51. Zheng, G. X. Y. et al. Massively parallel digital transcriptional profiling of single cells. *Nat. Commun.* **8**, 14049 (2017).
52. *R: A Language and Environment for Statistical Computing* (R Core Team, 2022).
53. Young, M. D. & Behjati, S. SoupX removes ambient RNA contamination from droplet-based single-cell RNA sequencing data. *Gigascience* **9**, giaa151 (2020).
54. Hippen, A. A. et al. miQC: an adaptive probabilistic framework for quality control of single-cell RNA-sequencing data. *PLoS Comput. Biol.* **17**, e1009290 (2021).
55. Human T cells from a healthy donor, 1k cells – multi (v2). *10X Genomics* https://www.10xgenomics.com/datasets/human-t-cells-from-a-healthy-donor-1-k-cells-multi-v-2-2-standard-5-0-0 (2020).
56. Szabo, P. A. et al. Single-cell transcriptomics of human T cells reveals tissue and activation signatures in health and disease. *Nat. Commun.* **10**, 4706 (2019).
57. Gao, S. et al. Single-cell RNA sequencing coupled to TCR profiling of large granular lymphocyte leukemia T cells. *Nat. Commun.* **13**, 1982 (2022).
58. Hao, Y. et al. Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573–3587.e29 (2021).

59. Head, T. et al. scikit-optimize/scikit-optimize (v0.9.0). *Zenodo* https://doi.org/10.5281/zenodo.5565057 (2021).
60. Andreatta, M. & Carmona, S. J. UCell: robust and scalable single-cell gene signature scoring. *Comput. Struct. Biotechnol. J.* **19**, 3796–3798 (2021).
61. Berset, M. et al. Expression of Melan-A/MART-1 antigen as a prognostic factor in primary cutaneous melanoma. *Int. J. Cancer* **95**, 73–77 (2001).

Acknowledgements

We gratefully acknowledge the data storage service SDS@hd supported by the Ministry of Science, Research and the Arts Baden-Württemberg and the German Research Foundation through grant INST 35/1503-1 FUGG, as well as the German Cancer Research Center Next Generation Sequencing and Flow Cytometry Core Facilities. This work was supported by grants from the TCR-POC program of the National Center for Tumor Diseases Heidelberg to M.P., I.P. and E.W.G., the Deutsche Forschungsgemeinschaft (German Research Foundation) project ID numbers 404521405; SFB1389—UNITE Glioblastoma, WP B01, project ID 394046768; The German Ministry of Education and Science National Center for Tumor Diseases Heidelberg 3.0 and German Cancer Consortium program 'Precision immunotherapy of brain tumors', the flagship 'Engineering Molecular Systems' spotlight proposal 'Synthetic Immunology', the kick-start seed funding project ENIGMA from the Hector Foundation and the CLINNOVA project 'Unlocking the potential of data science and AI in health care' by the Ministry of Science, Research and Arts Baden-Württemberg to M.P., the Dr. Rolf M. Schwiete Foundation project 'T cell receptor transgenic T cell therapy for patients with brain tumors' to M.P. and L.B., and the German Cancer Consortium joint funding UPGRADE program AMI2GO to L.B. and M.P. The K.H. Bauer Foundation supported R.O. and I.P. The Chinese Scholarship Council supported Z.M. The Sino-German Laboratory of Personalized Medicine for Pancreatic Cancer supported Z.M. and R.O. The Helmholtz Institute for Translational Oncology supported A.R.E. and T.B. Parts of Figs. 1–3 were created with BioRender.com.

Author contributions

C.L.T., K.L., T.B., Z.M., A.R.E., M.R., J.M.L. and G.H. carried out TCR testing experiments and analyzed results. C.L.T., Z.M. and E.W.G. performed bioinformatic analyses. M.R. performed surgery. K.L., L.B. and I.P. managed clinical trial sample acquisition, storage and analysis. A.D.R., A.F., F.I. and R.P.H. performed tumor killing assays. C.L.T., E.W.G., L.B., R.O. and M.P. conceived the work. E.W.G., C.L.T. and M.P. wrote the paper.

Funding

Open access funding provided by Deutsches Krebsforschungszentrum (DKFZ).

Competing interests

C.L.T., K.L., Z.M., A.R.E., R.O., L.B., M.P. and E.W.G. are inventors on patent applications describing the identification of tumor-reactive TCRs. M.P., R.P.H., R.O. and E.W.G. are founders of Tcelltech GmbH. A.F. and F.I. are employees of Cellply Srl. J.M.L. is an employee of BioMed X GmbH. The other authors declare no competing interests.

Additional information

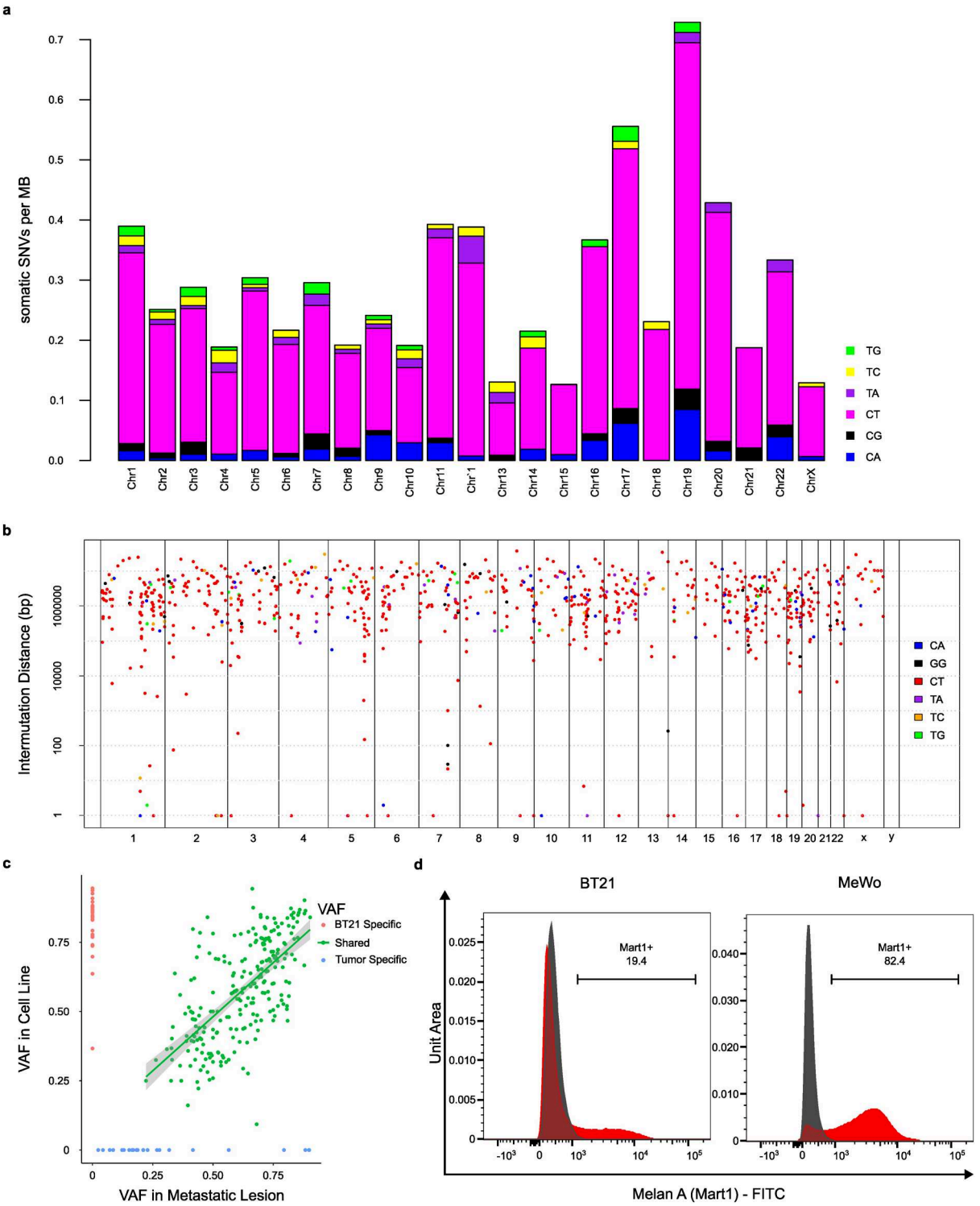
Extended data is available for this paper at https://doi.org/10.1038/s41587-024-02161-y.

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41587-024-02161-y.

Correspondence and requests for materials should be addressed to M. Platten or E. W. Green.

Peer review information *Nature Biotechnology* thanks Samra Turajlic and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

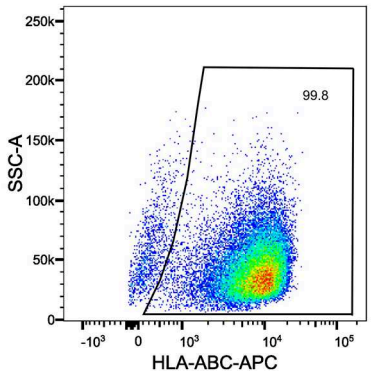
https://doi.org/10.1038/s41587-024-02161-y



Extended Data Fig. 1 | BT21 displays a pattern of mutations characteristic of melanoma. a: Metastatic melanoma sample BT21 displays the predominance of UV mediated C > T transition mutations typical of melanoma. **b**: Visualisation of mutations across chromosomes and their density, showing mutations spread throughout the genome, again typical of primary melanoma. **c**: Variant allele frequency (VAF) diagram showing overlap between the SNVs found in BT21 metastatic tumor sample and the BT21 cell line; the cell line's additional mutations may be the result of tumor heterogeneity or arose during the process of adaptation to cell culture conditions. Full SNV data in *.vcf format are included in Source Data 1. The error bands represent standard error of residuals. **d**: Flow cytometric analysis confirms expression of the tumor associated antigen MART1 in the BT21 cell line as found in 90% of primary cutaneous melanomas[61]. The commonly used MeWo melanoma cell line is shown for comparison.

https://doi.org/10.1038/s41587-024-02161-y

a



b

BT21 Tumor

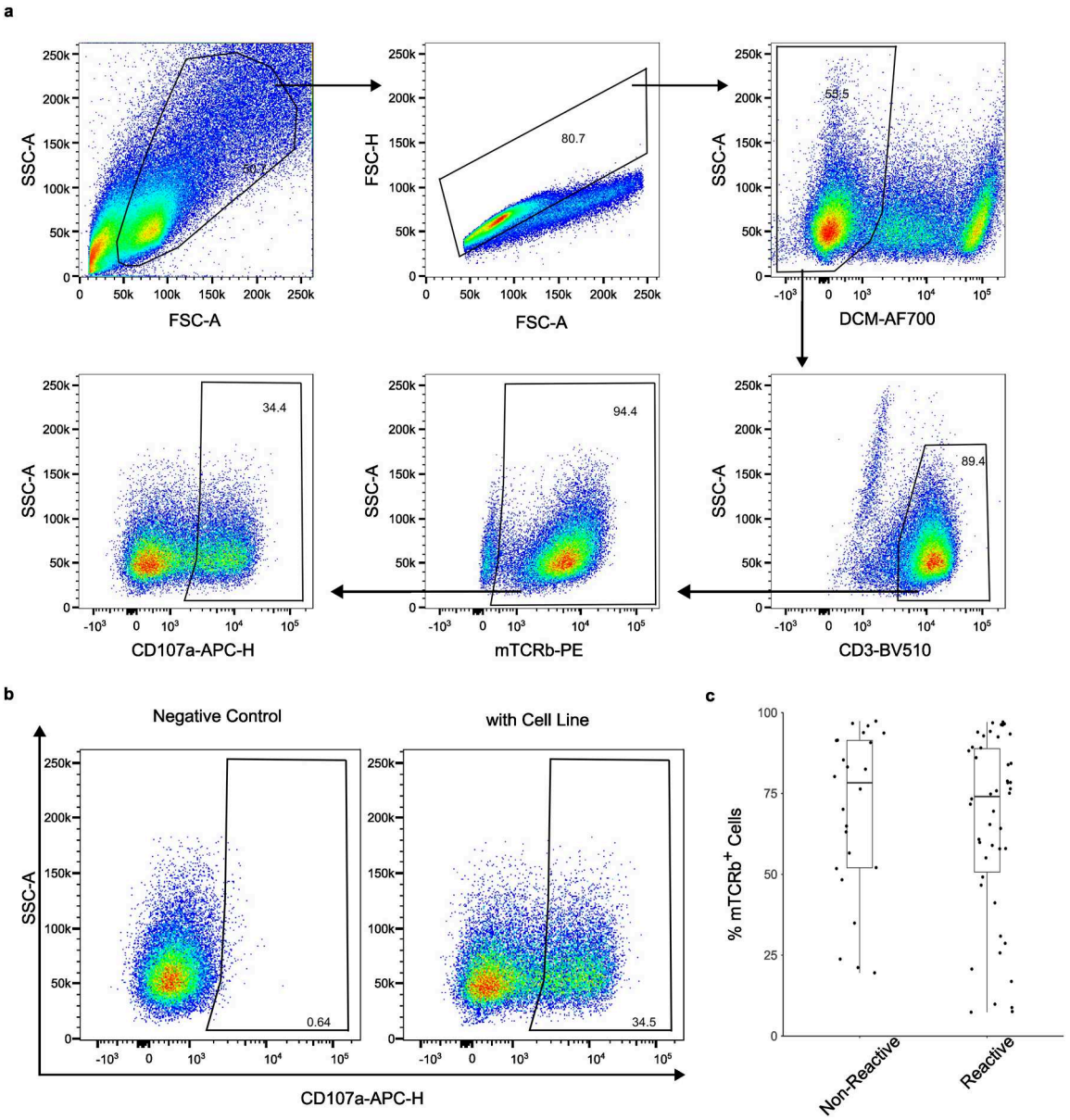
Gene	Alleles	Abundance	Read Count
HLA-A	A*24:02:01 / A*24:02:01	17.89%	61096
HLA-B	B*15:01:01 / B*07:02:01	2.52%	8527
HLA-C	C*03:03:01 / C*07:02:01	4.38%	14990

BT21 Cell Line

Gene	Alleles	Abundance	Read Count
HLA-A	A*24:02:01 / A*24:02:01	7.82%	14731
HLA-B	B*15:01:01 / B*07:02:01	2.71%	5064
HLA-C	C*03:03:01 / C*07:02:01	4.22%	7973

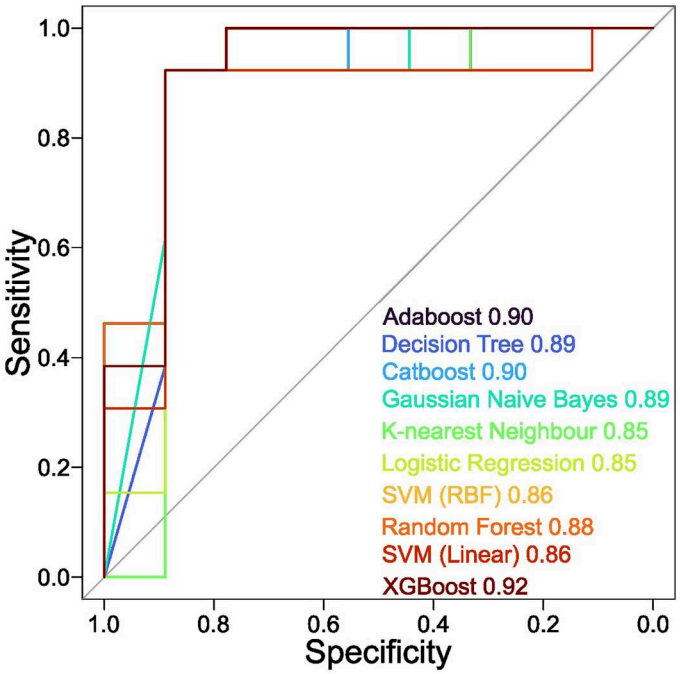
Extended Data Fig. 2 | BT21 cell line is suitable for immune testing. a: BT21 cell line expresses strong levels of MHC I after two days of stimulation with 300 IU/mL IFNγ. **b**: in silico HLA typing using arcasHLA[43] confirms that the BT21 tumor cell line is derived from primary tumor, and that HLA-A, -B and -C genes are expressed in the absence of stimulation. The 'Abundance' metric expresses the percentage of RNAseq reads mapping to HLA alleles, and gives a relative measure of expression between HLA loci. The 'Read Count' expresses the absolute number of RNAseq reads mapping to the alleles.

https://doi.org/10.1038/s41587-024-02161-y



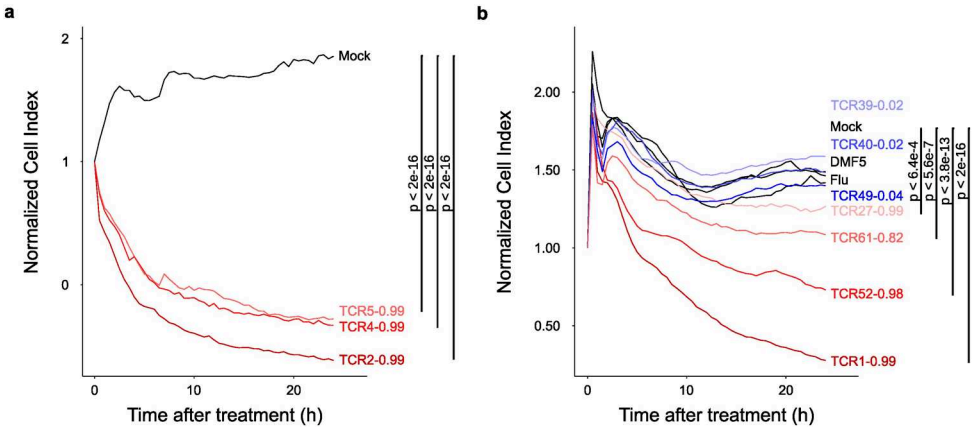
Extended Data Fig. 3 | TIL derived TCR testing data using cell line BT21. **a**: Gating strategy for TCR testing. It was gated for lymphocytes, singlet followed by live cells, then $CD3^+$ for T Cell, $mTCRb^+$ for transgenic T cells then subsequently CD107a for reactivity. **b**: Exemplary flow cytometry data showing the quantification of CD107a of the most reactive BT21 TCR with and without co-culture with BT21 cell line. After co-culture, 34.5% of cells were positive for CD107a. **c**: No significant difference in membrane expression of tumor-reactive and nonreactive TCR clonotypes. The centre line represents median, the lower and upper hinges corresponds to the first and third quantiles, and the upper whisker extends to maxima and lower whisker extends to minima (n = 26 for nonreactive TCR, n = 46 for reactive TCR).

https://doi.org/10.1038/s41587-024-02161-y



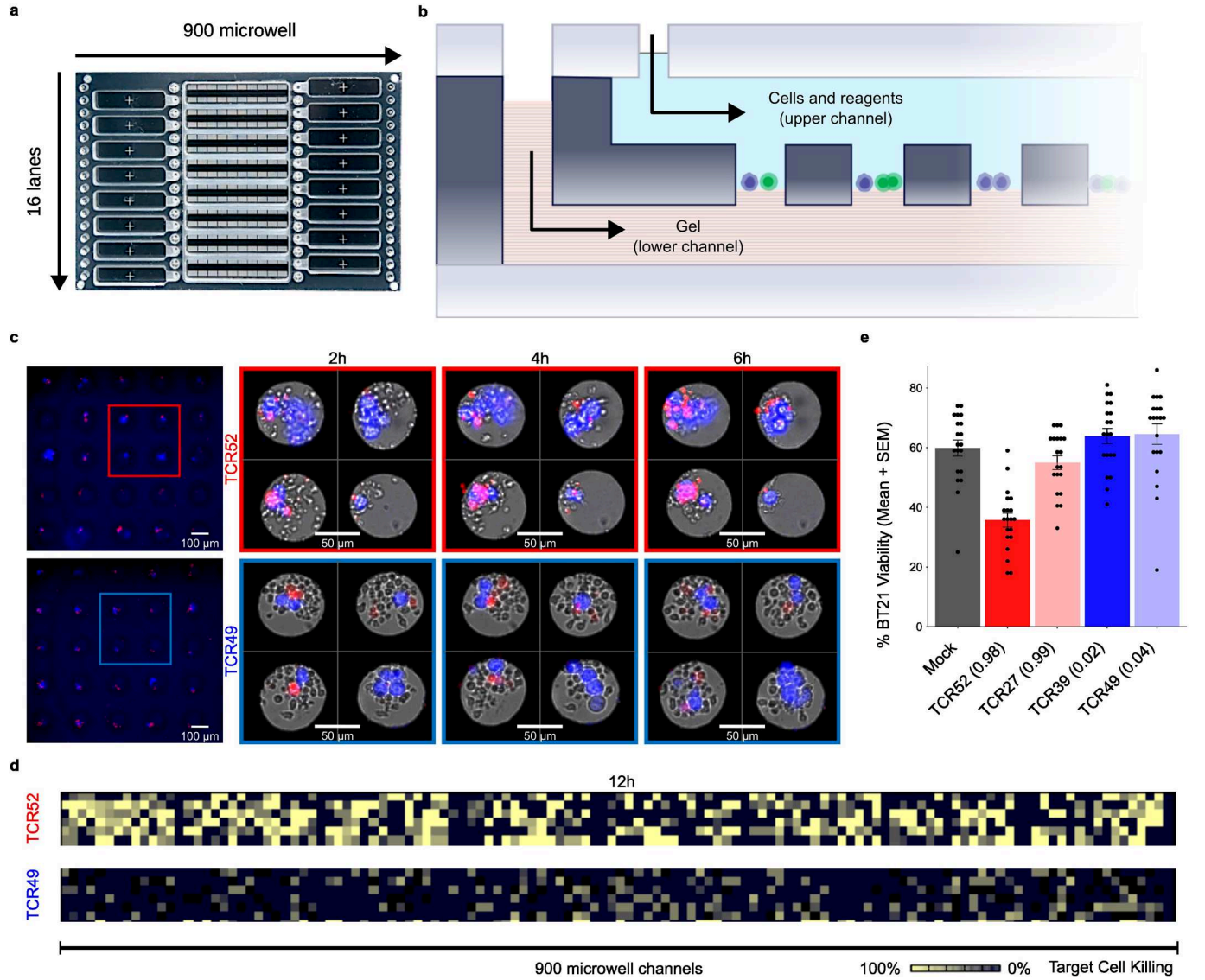
Extended Data Fig. 4 | Evaluation of the performance of different machine learning frameworks for TCR tumor-reactivity prediction. All models were implemented using the scikit-learn package, with the exception of XGBoost from xgboost package. The AUC represents model performance without further hyperparameter tuning. Models were trained using the data from 50 BT21 TCRs.

https://doi.org/10.1038/s41587-024-02161-y



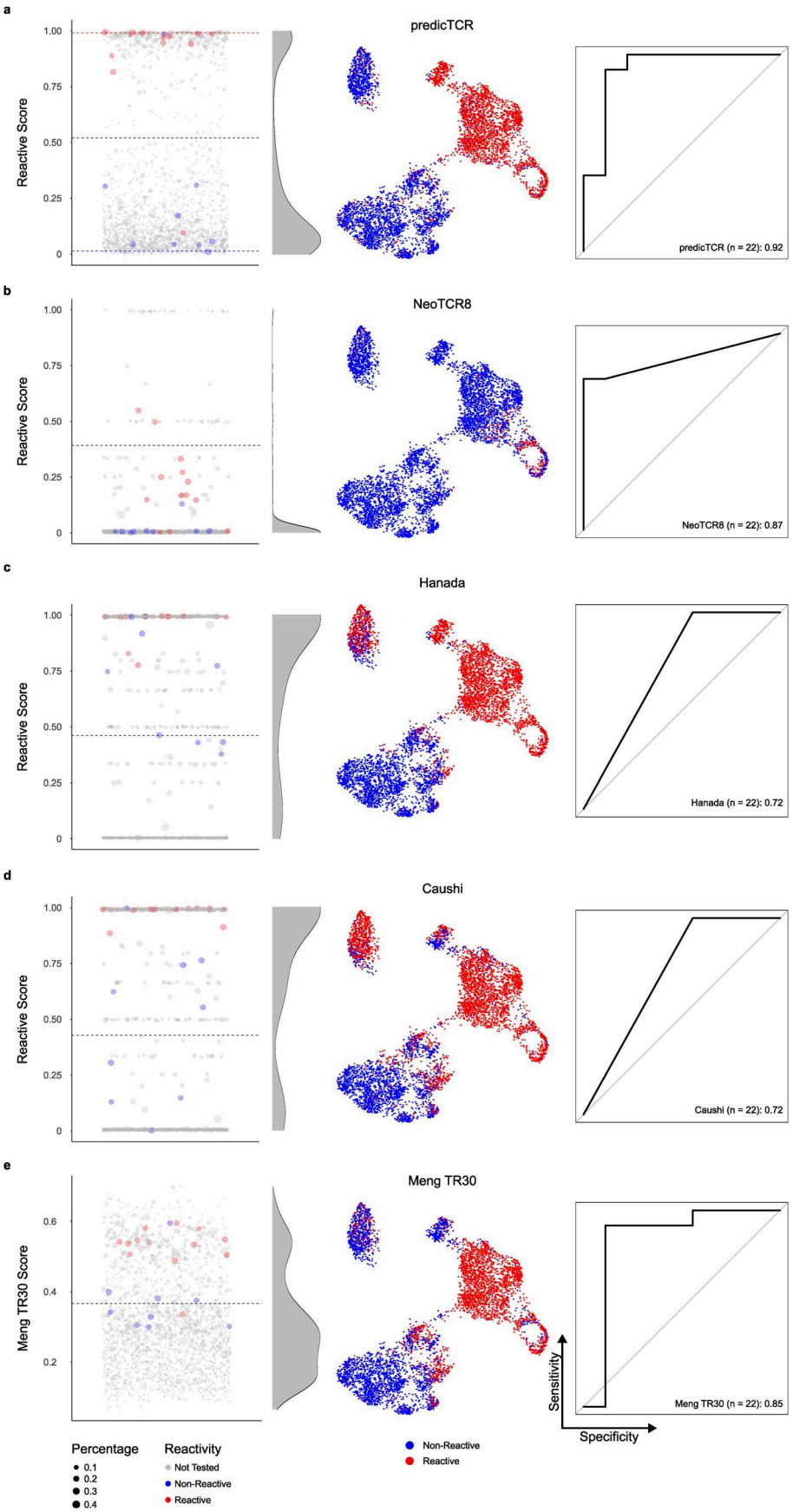
Extended Data Fig. 5 | Selected BT21 TIL derived TCRs can kill the BT21 cell line in co-culture assays. a: xCELLigence assay showing healthy donor PBMCs transfected with CD8+ TCRs that show high levels of CD107a when co-cultured with BT21 cells (Fig. 1d) can kill BT21 cells in a co-culture assay. **b**: Additional co-culture experiment testing killing capacity of BT21 derived TCRs, including TCRs that predicTCR50 designated as tumor-reactive. Assay was performed using PBMCs from the same healthy donor as in a. Lines are the average of replicates. Statistical analysis was performed as a one-way ANOVA on endpoint normalized cell index with a post-hoc Dunnett test comparison to the Mock. The predicTCR50 score for each TCR clonotype was calculated and added to the plot.

https://doi.org/10.1038/s41587-024-02161-y



Extended Data Fig. 6 | Single cell analysis of killing capacity of TCRs identified by predicTCR. a: Micrograph of the CC-Array microfluidic chip measured by the VivaCyteTM platform. **b**: Each lane of the CC-Array device comprises a lower microfluidic channel where the hydrogel is loaded, an upper microfluidic channel where cells and reagents are delivered and an array of microwells opened to both the upper and lower channels. Once the cells are loaded in the upper channel, they settle at the interface with the gel; random loading results in a distribution of effector:tumor cell ratios across each microfluidic channel. **c**: Fluorescence images of a subset of microwells containing a co-culture of PBMCs expressing a transgenic TCR and BT21 target cells at three timepoints (each timepoint shows 4 microwells drawn from 900 replicate microwells). The increase of PI signal (red) on CMA-stained cells (blue) over time reflects killing of the BT21 target cells by the transgenic effector T cells. **d:** Each well is quantified by the VivaCyte software to produce an overview of the killing activity of each TCR. **e**: Quantification of results for four different TCRs recapitulates the xCELLigence data (Extended Data Fig. 5b), with TCR52 performing more efficient target killing than TCR 27 after 12 hours co-culture. predicTCR reactivity scores for each TCR shown in parenthesis. Data are presented as mean values ± SEM for a single experimental run, with each mean representing the average of 20 fields of view (FOV), with each FOV consisting of an average of 45 microwells.

https://doi.org/10.1038/s41587-024-02161-y

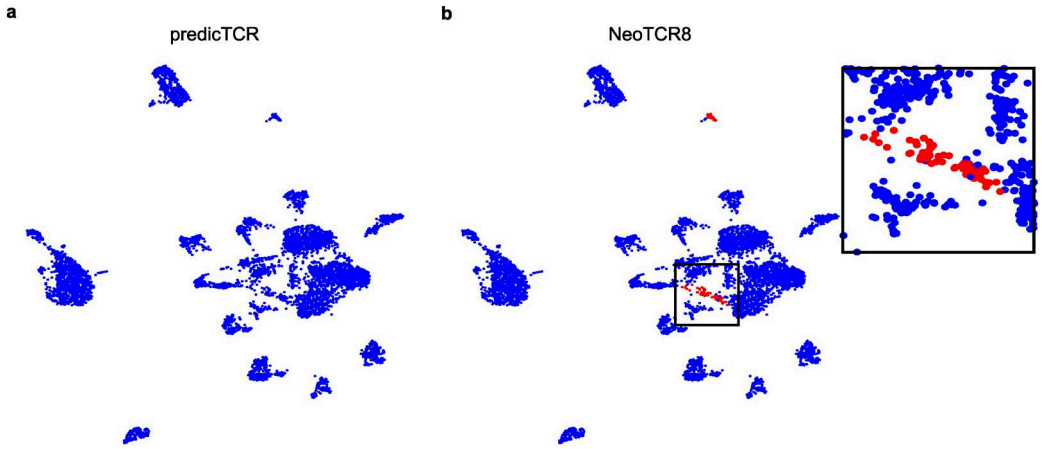


Extended Data Fig. 7 | See next page for caption.

https://doi.org/10.1038/s41587-024-02161-y

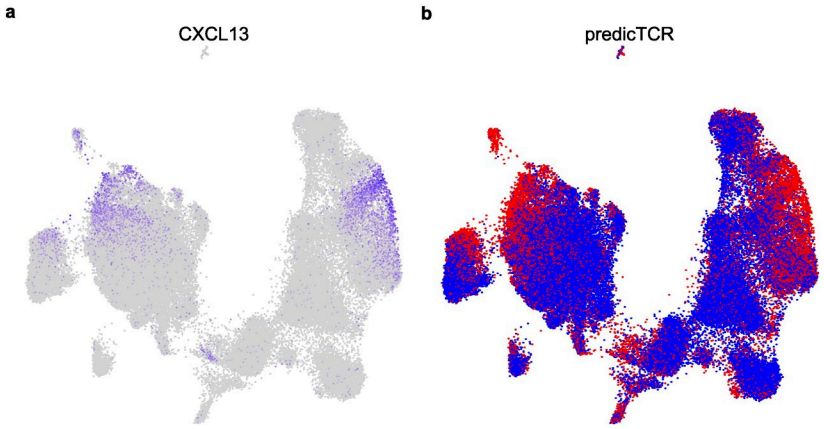
Extended Data Fig. 7 | Evaluation of overall performance of gene signature-based TCR tumor-reactivity predictions on BT21 derived TCRs. The left panel in each row shows the distribution of mean tumor-reactivity prediction scores for TIL derived TCR clonotypes from BT21, with a grey shaded density plot on the right. Coloured points show TCR clonotypes in the BT21 validation set (n = 22). TCRs are coloured by experimentally validated reactivity, with the size of each point reflecting the frequency of that TCR clonotype in the TILs. The grey dotted line indicates the sample-specific natural break threshold calculated to separate tumor-reactive from nonreactive clonotypes, while the red and blue dotted lines represent the 95th-percentile and 5th-percentile of tumor-reactivity scores respectively in panel (a). The central panel plots per-cell predictions onto the UMAP plot as in Fig. 1b. The right panel shows the ROC curve and AUC value for each classifier, for details see Table 1. **a**: performance of predicTCR50, showing a clear bimodal distribution of scores that correctly predicts the reactivity of 20 of the 22 validation set TCRs. Note that the false positive prediction (blue dot near top of plot) lies below the 95% threshold line. **b**: Performance of NeoTCR8 from Lowery et al. showing poor prediction of tumor-reactivity for the 22 TCR test set. **c**: Performance of signature developed by Hanada et. al. showing a high false-positive detection rate; note prediction of tumor-reactive T cells in the top left cluster (regulatory T cells, see Fig. 1b). **d**: Performance of signature developed by Caushi et al., also showing a similarly high false positive rate on BT21 TILs. **e**: Performance of signature TR30 from Meng et al. showing higher false positive rate than predicTCR50.

https://doi.org/10.1038/s41587-024-02161-y



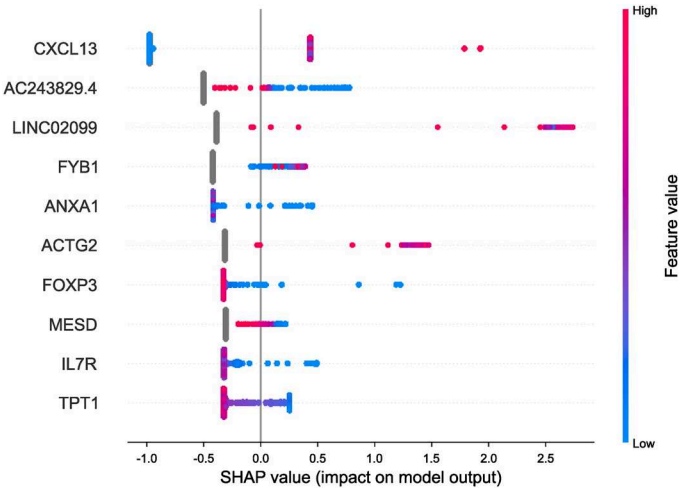
Extended Data Fig. 8 | UMAP showing performance of predicTCR and NeoTCR8 on PBMCs from a patient (PP4) with a severely symptomatic COVID-19 infection. a: Performance of predicTCR on PBMCs from patient PP4, with no cells predicted to be tumor-reactive. **b**: Performance of NeoTCR8 on same patient; inset panel shows a magnified view of a cluster containing T cells (n = 130, 2% of total cells) called to be tumor-reactive. PBMC data was obtained from Yoshida et al.[28].

https://doi.org/10.1038/s41587-024-02161-y



Extended Data Fig. 9 | UMAP showing that expression of CXCL13 does not always determine a defined cluster of TILs, but is similar to predicTCR predictions. a: T cells expressing CXCL13 in MD01-005 patient do not form a distinct cluster following dimension reduction; this reinforces the difficulty of obtaining accurate results from cluster-based gene set enrichment analyses. **b:** predicTCR tumor-reactivity predictions (which also rely in part on CXCL13) are not limited to a single cluster, and include (but are not limited to) the CXCL13 expressing cells.

https://doi.org/10.1038/s41587-024-02161-y



Extended Data Fig. 10 | Explainable AI SHAP on predicTCR. SHAP was performed on the predicTCR classifier to identify the top features (genes) that contribute to the classifier. The top 10 genes are illustrated as a beeswarm plot, with each dot representing a single cell. The colour indicates the expression value of a particular gene, and the x-axis indicate the SHAP value, the impact a particular feature on the model output. A positive SHAP value contribute to reactive prediction and vice versa.

nature portfolio

Corresponding author(s): Michael Platten

Last updated by author(s): 10.01.24

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐ A description of all covariates tested
☒	☐ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Single Cell Sequencing: 10X Chromium Single Cell V(D)J Reagent Kits (v1.1), sequenced on Novaseq 6000 (Illumina) Flow Cytomery: Samples were acquired using BD FACSLyrics device and BD FACS Suite (v1.5)
Data analysis	Flow Cytometry: FlowJo v10.6.2 Single Cell Sequencing: Alignment with cellranger v6.1.2, data normalization and transformation using SCTransform from Seurat (v4.3.0) General data analysis were done in R (v4.2.3) or Python (v3.9.16), with tidyverse (v2.0.0) or numpy (v1.24.3) or pandas (v2.0.2). All additional softwares and packages used are detailed in the Methods sections including: arcasHLA (v0.5.0), TRUST4 (v1.0.9), scikit-learn (v1.2.2), scikit-optimize (v0.9.0), shap (v0.41.0), GSVA (v1.46.0), pROC(1.18.2), caret (v6.0-94), xgboost (v1.7.4)

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

Single cell sequencing data for BT21 have been deposited in NCBI BioProject with accession code PRJNA985415. External datasets were obtained from GSE123139, GSE173351 and phs002748.v1.p1.

Human research participants

Policy information about studies involving human research participants and Sex and Gender in Research.

Reporting on sex and gender	The BT21 cell line was derived from a male patient. The sex of patients from external datasets is reported in each respective paper.
Population characteristics	62 year old male patient with a brain metastasis of a previously diagnosed with melanoma.
Recruitment	The patient was recruited in University Hospital Mannheim and provided written consent. We were able to able to obtain very fresh resection material which contributed to our successful establishment of the BT21 cell line.
Ethics oversight	Study was approved by the institutional review board (Ethikkommission 2019-643N)

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No sample size calculation was performed. The most frequent TCRs were chosen for model training, and the number of TCRs tested for a single sample is of a similar size or larger than in previous studies. The validation sample size was based on data availability of publicly available datasets.
Data exclusions	Low quality single cells are excluded from further analysis as described in Methods for failing to meet quality control thresholds. External datasets where no TCR validation done were excluded and patients with either only reactive or non-reactive TCR were omitted.
Replication	The microwell data in Extended Data Figures 5 and 6 replicates the reactivity testing data in Fig.1, confirming reactivity of these TCRs against the BT21 cell line.
Randomization	Random train/test set splits were used when training models in order to evaluate them, as common practice.
Blinding	The investigators performing the experimental TCR testing analyses (FACS, Cellply) were blinded as to the expected reactivity of TCRs tested.

Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	*Briefly describe the study type including whether data are quantitative, qualitative, or mixed-methods (e.g. qualitative cross-sectional, quantitative experimental, mixed-methods case study).*
Research sample	*State the research sample (e.g. Harvard university undergraduates, villagers in rural India) and provide relevant demographic*

Research sample	*information (e.g. age, sex) and indicate whether the sample is representative. Provide a rationale for the study sample chosen. For studies involving existing datasets, please describe the dataset and source.*
Sampling strategy	*Describe the sampling procedure (e.g. random, snowball, stratified, convenience). Describe the statistical methods that were used to predetermine sample size OR if no sample-size calculation was performed, describe how sample sizes were chosen and provide a rationale for why these sample sizes are sufficient. For qualitative data, please indicate whether data saturation was considered, and what criteria were used to decide that no further sampling was needed.*
Data collection	*Provide details about the data collection procedure, including the instruments or devices used to record the data (e.g. pen and paper, computer, eye tracker, video or audio equipment) whether anyone was present besides the participant(s) and the researcher, and whether the researcher was blind to experimental condition and/or the study hypothesis during data collection.*
Timing	*Indicate the start and stop dates of data collection. If there is a gap between collection periods, state the dates for each sample cohort.*
Data exclusions	*If no data were excluded from the analyses, state so OR if data were excluded, provide the exact number of exclusions and the rationale behind them, indicating whether exclusion criteria were pre-established.*
Non-participation	*State how many participants dropped out/declined participation and the reason(s) given OR provide response rate OR state that no participants dropped out/declined participation.*
Randomization	*If participants were not allocated into experimental groups, state so OR describe how participants were allocated to groups, and if allocation was not random, describe how covariates were controlled.*

Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	*Briefly describe the study. For quantitative data include treatment factors and interactions, design structure (e.g. factorial, nested, hierarchical), nature and number of experimental units and replicates.*
Research sample	*Describe the research sample (e.g. a group of tagged Passer domesticus, all Stenocereus thurberi within Organ Pipe Cactus National Monument), and provide a rationale for the sample choice. When relevant, describe the organism taxa, source, sex, age range and any manipulations. State what population the sample is meant to represent when applicable. For studies involving existing datasets, describe the data and its source.*
Sampling strategy	*Note the sampling procedure. Describe the statistical methods that were used to predetermine sample size OR if no sample-size calculation was performed, describe how sample sizes were chosen and provide a rationale for why these sample sizes are sufficient.*
Data collection	*Describe the data collection procedure, including who recorded the data and how.*
Timing and spatial scale	*Indicate the start and stop dates of data collection, noting the frequency and periodicity of sampling and providing a rationale for these choices. If there is a gap between collection periods, state the dates for each sample cohort. Specify the spatial scale from which the data are taken*
Data exclusions	*If no data were excluded from the analyses, state so OR if data were excluded, describe the exclusions and the rationale behind them, indicating whether exclusion criteria were pre-established.*
Reproducibility	*Describe the measures taken to verify the reproducibility of experimental findings. For each experiment, note whether any attempts to repeat the experiment failed OR state that all attempts to repeat the experiment were successful.*
Randomization	*Describe how samples/organisms/participants were allocated into groups. If allocation was not random, describe how covariates were controlled. If this is not relevant to your study, explain why.*
Blinding	*Describe the extent of blinding used during data acquisition and analysis. If blinding was not possible, describe why OR explain why blinding was not relevant to your study.*

Did the study involve field work? ☐ Yes ☐ No

Field work, collection and transport

Field conditions	*Describe the study conditions for field work, providing relevant parameters (e.g. temperature, rainfall).*
Location	*State the location of the sampling or experiment, providing relevant parameters (e.g. latitude and longitude, elevation, water depth).*
Access & import/export	*Describe the efforts you have made to access habitats and to collect and import/export your samples in a responsible manner and in*

Access & import/export	*compliance with local, national and international laws, noting any permits that were obtained (give the name of the issuing authority, the date of issue, and any identifying information).*
Disturbance	*Describe any disturbance caused by the study and how it was minimized.*

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☐	☒ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☐	☒ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	anti-CD3-BV510 (BD, cat 564713, clone HIT3A), anti-mTCRbeta-PE (BioLegend UK, cat 109208, clone H57-597), anti-CD107a-APC-H7 (BD, cat 561343, clone H4A3), anti-TNF-alpha-BV711 (Biolegend, cat 502940, clone MAb11), anti-Melan A-FITC (Santa Cruz Technology, cat sc-20032, clone A103)
Validation	All antibodies were used according to manufacturer's instruction or titrated prior to use (dilutions listed in paper methods section). anti-CD3-BV510: Manufacturer's quality tested for flow cytometry anti-mTCRbeta-PE: Manufacturer's quality tested for flow cytometry anti-CD107a-APC-H7: Manufacturer's quality tested for intracellular flow cytometry staining anti-TNF-alpha-BV711: Manufacturer's quality tested for intracellular flow cytometry staining

Eukaryotic cell lines

Policy information about cell lines and Sex and Gender in Research

Cell line source(s)	BT21: a cell line derived in-house from a male patient
Authentication	Exome-sequencing was done tumor material and the BT21 cell line material
Mycoplasma contamination	All cell lines tested negative for mycoplasma contamination
Commonly misidentified lines (See ICLAC register)	No commonly misidentified cell lines were used in this study

Palaeontology and Archaeology

Specimen provenance	*Provide provenance information for specimens and describe permits that were obtained for the work (including the name of the issuing authority, the date of issue, and any identifying information). Permits should encompass collection and, where applicable, export.*
Specimen deposition	*Indicate where the specimens have been deposited to permit free access by other researchers.*
Dating methods	*If new dates are provided, describe how they were obtained (e.g. collection, storage, sample pretreatment and measurement), where they were obtained (i.e. lab name), the calibration program and the protocol for quality assurance OR state that no new dates are provided.*

☐ Tick this box to confirm that the raw and calibrated dates are available in the paper or in Supplementary Information.

Ethics oversight	*Identify the organization(s) that approved or provided guidance on the study protocol, OR state that no ethical approval or guidance was required and explain why not.*

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Animals and other research organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research, and Sex and Gender in Research

Laboratory animals	*For laboratory animals, report species, strain and age OR state that the study did not involve laboratory animals.*
Wild animals	*Provide details on animals observed in or captured in the field; report species and age where possible. Describe how animals were caught and transported and what happened to captive animals after the study (if killed, explain why and describe method; if released, say where and when) OR state that the study did not involve wild animals.*
Reporting on sex	*Indicate if findings apply to only one sex; describe whether sex was considered in study design, methods used for assigning sex. Provide data disaggregated for sex where this information has been collected in the source data as appropriate; provide overall numbers in this Reporting Summary. Please state if this information has not been collected. Report sex-based analyses where performed, justify reasons for lack of sex-based analysis.*
Field-collected samples	*For laboratory work with field-collected samples, describe all relevant parameters such as housing, maintenance, temperature, photoperiod and end-of-experiment protocol OR state that the study did not involve samples collected from the field.*
Ethics oversight	*Identify the organization(s) that approved or provided guidance on the study protocol, OR state that no ethical approval or guidance was required and explain why not.*

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Clinical data

Policy information about clinical studies

All manuscripts should comply with the ICMJE guidelines for publication of clinical research and a completed CONSORT checklist must be included with all submissions.

Clinical trial registration	*Provide the trial registration number from ClinicalTrials.gov or an equivalent agency.*
Study protocol	*Note where the full trial protocol can be accessed OR if not available, explain why.*
Data collection	*Describe the settings and locales of data collection, noting the time periods of recruitment and data collection.*
Outcomes	*Describe how you pre-defined primary and secondary outcome measures and how you assessed these measures.*

Dual use research of concern

Policy information about dual use research of concern

Hazards

Could the accidental, deliberate or reckless misuse of agents or technologies generated in the work, or the application of information presented in the manuscript, pose a threat to:

No	Yes	
☒	☐	Public health
☒	☐	National security
☒	☐	Crops and/or livestock
☒	☐	Ecosystems
☒	☐	Any other significant area

Experiments of concern

Does the work involve any of these experiments of concern:

No	Yes	
☒	☐	Demonstrate how to render a vaccine ineffective
☒	☐	Confer resistance to therapeutically useful antibiotics or antiviral agents
☒	☐	Enhance the virulence of a pathogen or render a nonpathogen virulent
☒	☐	Increase transmissibility of a pathogen
☒	☐	Alter the host range of a pathogen
☒	☐	Enable evasion of diagnostic/detection modalities
☒	☐	Enable the weaponization of a biological agent or toxin
☒	☐	Any other potentially harmful combination of experiments and agents

ChIP-seq

Data deposition

☐ Confirm that both raw and final processed data have been deposited in a public database such as GEO.

☐ Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

Data access links *May remain private before publication.*	*For "Initial submission" or "Revised version" documents, provide reviewer access links. For your "Final submission" document, provide a link to the deposited data.*
Files in database submission	*Provide a list of all files available in the database submission.*
Genome browser session (e.g. UCSC)	*Provide a link to an anonymized genome browser session for "Initial submission" and "Revised version" documents only, to enable peer review. Write "no longer applicable" for "Final submission" documents.*

Methodology

Replicates	*Describe the experimental replicates, specifying number, type and replicate agreement.*
Sequencing depth	*Describe the sequencing depth for each experiment, providing the total number of reads, uniquely mapped reads, length of reads and whether they were paired- or single-end.*
Antibodies	*Describe the antibodies used for the ChIP-seq experiments; as applicable, provide supplier name, catalog number, clone name, and lot number.*
Peak calling parameters	*Specify the command line program and parameters used for read mapping and peak calling, including the ChIP, control and index files used.*
Data quality	*Describe the methods used to ensure data quality in full detail, including how many peaks are at FDR 5% and above 5-fold enrichment.*
Software	*Describe the software used to collect and analyze the ChIP-seq data. For custom code that has been deposited into a community repository, provide accession details.*

Flow Cytometry

Plots

Confirm that:

☒ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).

☒ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).

☒ All plots are contour plots with outliers or pseudocolor plots.

☒ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation	TCR-encoding RNA was electroporated into expanded healthy donor PBMCs using the Lonza 4D nucleofector (program EO-115, solution P3 supplemented according to the manufacturer's recommendations), which were plated into 48-well plates containing TexMACS media (Miltenyi 130-097-196) supplemented with 2% human AB serum. 18-24 hours after electroporation, cells were harvested and 50 IU/mL benzonase (Speed BioSystems YCP1200-50KU) were

added to avoid cell clumping. TCR expression levels were measured via flow cytometry with markers including fixable viability dye (AF700, BD), CD3 (clone HIT3A, BV510, BD), CD4 (clone SK3, BV786, BD), CD8 (clone RPA-T8, PerCP-Cy5.5, Invitrogen) and mTCRb (clone H57-597, PE, Biolegend).

To assess TCR reactivity, a total of 150,000 T cells and 75,000 cells of the patient-autologous tumor cell line were co-cultured in U-bottom 96-well plates in a total volume of 200 µL. Wells with only T cells, or T cells and TransAct beads (Miltenyi 130-111-160), were used as negative and positive controls, respectively. 5 µL of CD107a FACS antibody (BD, REF 561343) was added per well. After one hour of co-culture, GolgiPlug and GolgiStop (BD 555029 and 554724) were added to reach a 1:1000 dilution, and after four additional hours of co-culture cells were used for flow cytometry analysis.

Markers included fixable viability dye (AF700, eBioscience), CD3 (clone HIT3A, BV510, BD), CD4 (clone SK3, APC, BD), CD8 (clone RPA-T8, PerCP-Cy5.5, Invitrogen), mTCRb (clone H57-597, PE, Biolegend) and TNFa (clone MAb11, BV711, Biolegend). Samples were acquired on a FACS Lyric device, flow cytometry data was analysed using FlowJo software, v10.6.2 (FlowJo LLC).

Instrument: BD FACSLyrics, ZE5 Cell Analyzer

Software: BD FACS Suite (v1.5), FlowJo (v10.6.2)

Cell population abundance: Population abundance varies by TCR tested: the maximum amount of reactive (CD107a+) cells is 33%, ranging down to 0% for non-reactive TCRs. See example in Extended Data Figure 3b.

Gating strategy: Gated on lymphocytes based on size and granularity (FSC vs SSC) -> singlet -> live cells (negative on LIVE/DEAD) -> CD3+ -> mTCRb+ -> CD107a+ (Extended Data Figure 3a)

☒ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.

Magnetic resonance imaging

Experimental design

Design type: *Indicate task or resting state; event-related or block design.*

Design specifications: *Specify the number of blocks, trials or experimental units per session and/or subject, and specify the length of each trial or block (if trials are blocked) and interval between trials.*

Behavioral performance measures: *State number and/or type of variables recorded (e.g. correct button press, response time) and what statistics were used to establish that the subjects were performing the task as expected (e.g. mean, range, and/or standard deviation across subjects).*

Acquisition

Imaging type(s): *Specify: functional, structural, diffusion, perfusion.*

Field strength: *Specify in Tesla*

Sequence & imaging parameters: *Specify the pulse sequence type (gradient echo, spin echo, etc.), imaging type (EPI, spiral, etc.), field of view, matrix size, slice thickness, orientation and TE/TR/flip angle.*

Area of acquisition: *State whether a whole brain scan was used OR define the area of acquisition, describing how the region was determined.*

Diffusion MRI ☐ Used ☐ Not used

Preprocessing

Preprocessing software: *Provide detail on software version and revision number and on specific parameters (model/functions, brain extraction, segmentation, smoothing kernel size, etc.).*

Normalization: *If data were normalized/standardized, describe the approach(es): specify linear or non-linear and define image types used for transformation OR indicate that data were not normalized and explain rationale for lack of normalization.*

Normalization template: *Describe the template used for normalization/transformation, specifying subject space or group standardized space (e.g. original Talairach, MNI305, ICBM152) OR indicate that the data were not normalized.*

Noise and artifact removal: *Describe your procedure(s) for artifact and structured noise removal, specifying motion parameters, tissue signals and physiological signals (heart rate, respiration).*

Volume censoring: *Define your software and/or method and criteria for volume censoring, and state the extent of such censoring.*

Statistical modeling & inference

Model type and settings: *Specify type (mass univariate, multivariate, RSA, predictive, etc.) and describe essential details of the model at the first and second levels (e.g. fixed, random or mixed effects; drift or auto-correlation).*

Effect(s) tested

Define precise effect in terms of the task or stimulus conditions instead of psychological concepts and indicate whether ANOVA or factorial designs were used.

Specify type of analysis: ☐ Whole brain ☐ ROI-based ☐ Both

Statistic type for inference
(See Eklund et al. 2016)

Specify voxel-wise or cluster-wise and report all relevant parameters for cluster-wise methods.

Correction

Describe the type of correction and how it is obtained for multiple comparisons (e.g. FWE, FDR, permutation or Monte Carlo).

Models & analysis

n/a	Involved in the study
☐	☐ Functional and/or effective connectivity
☐	☐ Graph analysis
☐	☐ Multivariate modeling or predictive analysis

Functional and/or effective connectivity

Report the measures of dependence used and the model details (e.g. Pearson correlation, partial correlation, mutual information).

Graph analysis

Report the dependent variable and connectivity measure, specifying weighted graph or binarized graph, subject- or group-level, and the global and/or node summaries used (e.g. clustering coefficient, efficiency, etc.).

Multivariate modeling and predictive analysis

Specify independent variables, features extraction and dimension reduction, model, training and evaluation metrics.