

# Current Opinion in Immunology

## Applications of T cell receptor (TCR) repertoire analysis in the advancement of cancer immunotherapy. --Manuscript Draft--

<b>Short Title:</b>	TCR repertoire and cancer immunotherapy
<b>Keywords:</b>	TCR repertoire; diversity; T cell receptor; cancer immunotherapy
<b>Corresponding Author:</b>	Benny Chain UCL: University College London UNITED KINGDOM
<b>Corresponding Author's Institution:</b>	UCL: University College London
<b>Corresponding Author E-Mail:</b>	b.chain@ucl.ac.uk
<b>First Author:</b>	Kroopa Joshi
<b>Order of Authors:</b>	Kroopa Joshi Martina Milighetti Benny Chain
<b>Abstract:</b>	<p>T cell receptor sequencing is a powerful new approach to analysis of the host-tumour interaction. Advances in NextGen sequencing, coupled with powerful bioinformatic tools, allow quantitative reproducible characterisation of repertoires from tumour and blood from patients with a variety of cancers. We consider how global metrics such as T cell clonality and diversity be extracted from these repertoires give insight into the mechanism of action of immune checkpoint blockade. We explore how the analysis of TCR overlap between repertoires can help define spatial and temporal heterogeneity of the anti-tumoural immune response. Finally, we review how analysis of TCR sequence and structure can be used to annotate antigenic specificity, with important implications for the development of personalised adoptive cellular immunotherapies. Importantly, advancements in single cell TCR profiling, coupled with powerful analytical tools adopted from artificial intelligence have the potential to drive the next breakthrough in cancer immunotherapy.</p>
<b>Author Comments:</b>	<p>Thank you for reviewing our MS and for the helpful comments. We have edited the MS to address these points as follows:</p> <ol style="list-style-type: none"><li>1. On page 3 the authors talk about the observation that CTLA-4 blockage leads to increase TCR diversity. This requires further explanation; is it simply a consequence of local T cell recruitment (as suggested); does 'new recruitment' also apply to blood; are new TCR generated in the thymus? This is a very interesting point. Remarkably, the systemic effects of anti-CTLA4 on the overall repertoire seem to be remarkably limited, and the effects on intra-tumour heterogeneity more marked. Perhaps the concentrations achieved in vivo are such that the effect on the immune system as a whole (blood, thymus etc.) are relatively limited, and the effects are more marked where Tregs are actively limiting an immune response such as in a tumour. We have added a couple of sentences to explain this point on page 3</li><li>2. In the section 'Repertoire analysis in assigning TCR antigenic specificity' to authors should clarify whether cited papers use CDR3 motifs in one TCR chain, or combine CDR3 motifs in the alpha and beta chain. The authors should also clarify whether CDR1 and CDR2 are included in any of the specificity prediction models.  Although the prediction tools do not usually explicitly make use of the other CDRs, the structure-based ones implicitly make use of the whole TCR sequence. We have added a section to address this point at the top of page 7.</li><li>3. The left panel in figure 2B suggests that low mutation rates drive clonality of tumour infiltrating T cells. Is this the case? No, we do not wish to imply this; only that increased mutational heterogeneity leads to increased TCR heterogeneity. We have edited the figure somewhat, and added an explanatory sentence in the text (top of</li></ol>



## **Application of T cell receptor (TCR) repertoire analysis in the advancement of cancer immunotherapy.**

Kroopa Joshi<sup>1</sup>, Martina Milighetti<sup>2</sup>, Benjamin M Chain<sup>2,3</sup>

<sup>1</sup>. Department of Medical Oncology, The Royal Marsden NHS Foundation Trust, London, United Kingdom.

<sup>2</sup>. Division of Infection and Immunity, University College London, London, United Kingdom.

<sup>3</sup>. Department of Computer Science, University College London, London, United Kingdom.

### **Highlights**

- TCR repertoire analysis as a tool to advance cancer immunotherapy
- Computational methods are used to understand TCR repertoire characteristics
- Both TCR sequence and structure have been used for specificity prediction

### **Abstract**

T cell receptor (TCR) sequencing has emerged as a powerful new technology in analysis of the host-tumour interaction. The advances in NextGen sequencing technologies, coupled with powerful novel bioinformatic tools, allow quantitative and reproducible characterisation of repertoires from tumour and blood samples from an increasing number of patients with a variety of solid cancers. In this review, we consider how global metrics such as T cell clonality and diversity can be extracted from these repertoires and used to give insight into the mechanism of action of immune checkpoint blockade. Furthermore, we explore how the analysis of TCR overlap between repertoires can help define spatial and temporal heterogeneity of the anti-tumoural immune response. Finally, we review how analysis of TCR sequence and structure, either of individual TCRs or from sets of related TCRs can be used to annotate the antigenic specificity, with important implications for the development of personalised adoptive cellular immunotherapies. Importantly, advancements in single cell TCR profiling, coupled with powerful analytical tools adopted from artificial intelligence have the potential to drive the next breakthrough in cancer immunotherapy.

## **Introduction**

T cell receptor (TCR) repertoire sequencing is a powerful tool which can characterise the breadth, strength and dynamics of the anti-tumoural immune response. The advances in next generation (NextGen) sequencing technologies and associated bioinformatic pipelines have significantly improved the sensitivity, accuracy and quantitation of TCR repertoire studies, resulting in a large number of publications and associated data sets. However, major computational challenges remain in inferring the dynamics and antigen specificity from static snapshots of the TCR repertoire. Here, we review progress in applying TCR repertoire analysis to track the host immune responses to immunotherapeutic checkpoint blockade, to follow the evolution and heterogeneity of the response in the face of a changing tumour microenvironment and to reveal T cell antigen specificity.

## **T cell clonality and diversity as metrics for immune checkpoint TCR repertoire analysis**

TCRs consist of a heterodimer of alpha and beta, or gamma and delta chains. In this review we focus on alpha and beta TCRs. TCR genes are generated during T cell differentiation by stochastic and imprecise somatic cell recombination of sets of short homologous genomic sequences, the V, D and J genes. Because the total number of possible TCR alpha/beta heterodimers is huge<sup>1</sup> the same sequence is rarely generated twice in an individual, and so identical sequences observed multiple times are likely to be derived by selection and proliferation of a single T cell clone. The overall clonality of the TCR repertoire<sup>2</sup>, a metric which captures the number and frequency of TCRs which are observed multiple times in a sample, is often used as an indicator of immune activity (Figure 1). Repertoire diversity is inversely related to clonality, and captures the breadth of a repertoire.

Several groups have investigated the effect of immune checkpoint blockade on T cell clonality and diversity. Given the challenge of obtaining longitudinal tumour samples, the majority of these studies have utilised peripheral blood sampling. In patients with

bladder cancer, urine derived lymphocytes (UDLs), may offer an additional non-invasive method of TCR repertoire sequencing<sup>3</sup>.

### **Effects of anti-CTLA-4 on the TCR repertoire**

Increased TCR repertoire diversity following anti-CTLA-4 therapy is reported in patients with advanced melanoma.<sup>4</sup> This increased diversity has been hypothesised to reflect recruitment of new tumour-reactive clones, as suggested by an increase in the number of newly detected CD8<sup>+</sup> melanoma specific TCRs following anti-CTLA-4 treatment.<sup>5</sup> Surprisingly, the systemic effects on the TCR repertoire seem to be quite limited. Perhaps the concentrations achieved in vivo are such that the effect on the immune system as a whole (blood, thymus etc.) are relatively limited, and are limited to where Tregs are most active in limiting an immune response (such as in a tumour). However, increased peripheral blood T cell diversity has also been linked to anti-CTLA-4 toxicity, suggesting checkpoint inhibition may also result in recruitment of self-reactive T cells.<sup>6</sup> In contrast, increased repertoire clonality prior<sup>7</sup> to or following<sup>8</sup> treatment is associated with improved response to anti-CTLA-4 therapy and survival. Thus, TCR repertoire analysis supports the idea that strong pre-existing immunity, for example, increased T cell clonality, together with anti-CTLA-4 induced broadening of the T cell pool are key to deliver an effective immunotherapeutic response (Figure 2A).

### **Effects of anti-PD-1 on the TCR repertoire**

Similar to anti-CTLA-4, increased pre-treatment clonality has been associated with improved clinical outcomes to anti-PD1 therapy in melanoma<sup>9,10</sup>. Moreover, increased T cell clonality following anti-PD-1 treatment is associated with clinical response in metastatic melanoma, NSCLC, glioblastoma and metastatic bladder cancer.<sup>9,11,12,14,17,12</sup> Interestingly, increased T cell clonality detected in 'on-treatment' tumour samples was predominantly a feature of single agent anti-PD-1 treatment rather than those undergoing dual anti-PD-1/anti-CTLA-4 blockade. Accordingly, anti-PD-1 may favour the maintenance and expansion of existing anti-tumour T cells (Figure 2A). However, oligoclonal expansion of novel T cell clonotypes has been observed in basal and squamous cell carcinomas treated with anti-PD-1, suggesting that in some contexts pre-existing tumour infiltrating lymphocytes (TILs) may have

limited reinvigoration capacity and that the intratumoural T cell response to checkpoint blockade may be due to a distinct repertoire of T cell clones that have recently entered the tumour microenvironment.<sup>13</sup>

The functional heterogeneity of the T cell compartment in peripheral blood complicates the interpretation of the repertoire studies reported to date, and more detailed analysis of fractionated T cells using cell surface markers that enrich for anti-tumour T cells (e.g., CD8<sup>+</sup>PD1<sup>+</sup>) is an area of active research. Nevertheless, there is convincing evidence that the clonality and diversity of the TCR repertoire reflects the therapeutic response to checkpoint inhibitors, providing insight into mechanisms of action, and supporting the use of TCR repertoire analysis as a biomarker for stratification and monitoring of patients on immunotherapy.

### **TCR repertoire heterogeneity and tumour genomic heterogeneity**

Tumour heterogeneity, arising from genetic and epigenetic changes during tumour evolution, is a well-established and important feature of many cancers, and has a major influence on disease progression.<sup>14</sup> The relationship between tumour heterogeneity and the host immune response is complex, since the immune response is a powerful selection force driving immune evolution, but genomic instability can also produce neoantigens which drive immune responses with important clinical relevance. For example, studies have shown that a high burden of clonal nonsynonymous mutations is associated with reduced disease recurrence and improved response to checkpoint blockade<sup>15-17</sup>.

A key confounder in comparing TCR repertoires from distinct tumour samples is heterogeneity arising as an artefact of sampling during nucleic acid extraction, PCR amplification and sequencing. It is therefore important to establish a rigorous statistical framework to control for the sampling process.

In the context of early-stage NSCLC, we have developed such a statistically rigorous approach and used it to define ubiquitous TCRs, present throughout a tumour, and regional TCRs, found in only defined regions of a tumour.<sup>18</sup> Ubiquitous TCRs, present across multiple regions of the tumour or in multiple metastatic lesions may represent

responses to common tumour antigens found in all cancer cells, although direct experimental proof for this hypothesis is still limited. Since neoantigens can appear progressively during the course of tumour evolution, the number of neoantigens in different regions of the tumour may differ considerably. We have previously reported that different T cell clonotypes can also be observed in different regions of the tumour, and that spatial heterogeneity in expanded T cell clones is correlated with spatial mutational heterogeneity (illustrated schematically in Figure 2B). These findings, and numerous further reports of TCR heterogeneity across both primary<sup>19-21</sup> and metastatic tumours<sup>22-24</sup>, support the idea that expanded intra-tumoural T cell clones are driven by the intratumoural neoantigen landscape that may be sculpted by somatic mutations, and immune editing including focal HLA loss or antigen processing defects.<sup>15, 18</sup> Furthermore, heterogeneity of the repertoire in longitudinal samples may reflect dynamic changes in the host-tumour relationship.<sup>18</sup>

### **Repertoire analysis in assigning TCR antigenic specificity**

Assigning neoantigen or tumour-associated antigen specificity to TILs is an area of key interest, both to improve our understanding of the adaptive anti-tumoural immune response and also to translate this knowledge into therapy (Figure 2C). The identification of clusters or networks of sequence related TCRs within populations of TILs<sup>18</sup> provide strong circumstantial evidence of antigen specificity<sup>25</sup> (Figure 3A). TCR sequencing in combination with *in vitro* peptide-stimulation (mutation-associated neoantigen functional expansion of specific T cells (MANAFEST)<sup>26</sup> assay) and/or MHC-peptide multimer sorting<sup>18</sup> together with single cell RNA sequencing (scRNAseq) provide effective platforms with which to explore the antigen specificity of lymphocytes directly. In contrast to bulk sequencing, scRNAseq identifies alpha/beta gene pairs, and thus permits a much more complete characterisation of the TCR receptor associated with each T cell. In addition, combining TCR sequencing with global scRNAseq links the T cell receptor to the transcriptional state of T cells.<sup>27-29</sup> The majority of studies which have identified antigen-specific TCRs in this way have focused on TILs, since these are presumably highly enriched for tumour-reactive T cells. However, in one example TCR repertoire analysis demonstrated that approximately 11% of tumour resident T cell clones were detected in T cells expanded from PD-1 positive T cells in the peripheral blood.<sup>30</sup> Thus antigen-specificity studies

can also be carried out using the much more accessible pool of peripheral blood T cells.

## **Future directions**

Despite these elegant studies combining TCR repertoire sequencing with function to identify target antigens, the annotation of TCRs by functional approaches remains extremely challenging, labour intensive and costly, limiting its application to clinical studies outside of a few exceptionally well-funded research centres.

There is therefore intense interest in predicting pMHC antigen specificity of individual TCRs from sequence alone, an extremely challenging computational challenge. Several studies<sup>25, 31, 32</sup> have showed that TCRs that bind the same epitope show conserved sequence motifs, especially in the CDR3 loop, the TCR region that forms the most extensive contacts with the peptide epitope.<sup>33, 34</sup>

We and others have used various metrics of sequence similarity to develop TCR clustering algorithms to gain insight into the antigenic specificity of TCRs<sup>35-38</sup>. This is a powerful approach to identify related sequences in a repertoire, once the antigen specificity of some members of the cluster is known (Figure 3B). For example, the clustering algorithm GLIPH2 has recently been used to find TCR specificity groups in two independent NSCLC datasets, and in combination with experimental antigen discovery was able to identify T cells reacting to tumour antigens<sup>39</sup>.

More sophisticated machine-learning approaches to encoding the sequence information of available TCR-antigen pairs can also be used to predict TCR specificity (Figure 3C)<sup>25, 31, 40-43</sup>. Most available algorithms require training on existing datasets for prediction, i.e., a set of TCRs known to recognise a specific epitope is needed to be able to determine whether a new TCR will bind to that same antigen. However, the antigen driving the sequence convergence is often unknown, thus limiting the practical



application of these methods. To overcome this challenge, new classifiers that try and learn general binding rules using either sequence or structural<sup>44-46</sup> features have been created and applied to predict probability of binding for unknown TCRs and epitopes. Of note, different algorithms require different inputs: ERGO and ImRex only need the CDR3 $\beta$  and peptide sequences; GLIPH relies on the CDR3 $\beta$  sequence and V-gene usage; TCRdist shows the best performance when information about the three CDR loops (as well as a fourth loop between CDR2 and CDR3) is included from both chains. The structure-based classifiers, instead, base the prediction on the properties of the (predicted) conformation of the entire complex, which will depend on the sequence and structure of both TCR chains (including the CDR loops, but also the framework features that determine the loop positioning). As paired  $\alpha\beta$  chains are necessary for a structural approach, the sequence classifiers are more readily applicable to TCR-seq information, which does not provide pairing information. However, further work will be required to develop the accuracy of these computational approaches to a point at which they will be of immediate value to tumour immunology.

## **Conclusions**

TCR repertoire sequencing is increasingly becoming an indispensable part of the cancer immunologist's arsenal. As sequencing costs fall further, protocols become more standardised, and computational tools more reliable and user-friendly, repertoire measurements will increasingly find a use in a clinical context, for monitoring of disease recurrence and progression, assessment of response to systemic treatment and patient stratification. However, perhaps, the most powerful impact of repertoire analysis will emerge from the application of the latest advances in artificial intelligence which will ultimately allow annotation of the repertoire in terms of antigen specificity of the anti-tumoural immune response with important clinical relevance for the development of adoptive cellular therapies for patients with solid tumours.

## **Declaration of interest**

None

## **Acknowledgements**

The authors gratefully acknowledge grant support from Cancer Research UK, The Rosetrees Foundation and The National Institute for Health Research UCL Hospitals Biomedical Research Centre.

## References

1. Elhanati Y, Murugan A, Callan CG, Mora T, Walczak AM. Quantifying selection in immune receptor repertoires. *Proc Natl Acad Sci U S A*. Jul 2014;111(27):9875-80. doi:10.1073/pnas.1409572111
2. Mahe E, Pugh T, Kamel-Reid S. T cell clonality assessment: past, present and future. *J Clin Pathol*. Mar 2018;71(3):195-200. doi:10.1136/jclinpath-2017-204761
3. Wong YNS, Joshi K, Khetrpal P, et al. Urine-derived lymphocytes as a non-invasive measure of the bladder tumor immune microenvironment. *J Exp Med*. Sep 2018;doi:10.1084/jem.20181003
4. Robert L, Tsoi J, Wang X, et al. CTLA4 blockade broadens the peripheral T-cell receptor repertoire. *Clin Cancer Res*. May 2014;20(9):2424-32. doi:10.1158/1078-0432.CCR-13-2648
5. Kvistborg P, Philips D, Kelderman S, et al. Anti-CTLA-4 therapy broadens the melanoma-reactive CD8+ T cell response. *Sci Transl Med*. Sep 2014;6(254):254ra128. doi:10.1126/scitranslmed.3008918
6. Oh DY, Cham J, Zhang L, et al. Immune Toxicities Elicited by CTLA-4 Blockade in Cancer Patients Are Associated with Early Diversification of the T-cell Repertoire. *Cancer Res*. 03 2017;77(6):1322-1330. doi:10.1158/0008-5472.CAN-16-2324
7. Cha E, Klinger M, Hou Y, et al. Improved survival with T cell clonotype stability after anti-CTLA-4 treatment in cancer patients. *Sci Transl Med*. May 2014;6(238):238ra70. doi:10.1126/scitranslmed.3008211
8. Hopkins AC, Yarchoan M, Durham JN, et al. T cell receptor repertoire features associated with survival in immunotherapy-treated pancreatic ductal adenocarcinoma. *JCI Insight*. Jul 2018;3(13)doi:10.1172/jci.insight.122092
9. Tumeh PC, Harview CL, Yearley JH, et al. PD-1 blockade induces responses by inhibiting adaptive immune resistance. *Nature*. Nov 2014;515(7528):568-71. doi:10.1038/nature13954
10. Hogan SA, Courtier A, Cheng PF, et al. Peripheral Blood TCR Repertoire Profiling May Facilitate Patient Stratification for Immunotherapy against Melanoma. *Cancer Immunol Res*. Jan 2019;7(1):77-85. doi:10.1158/2326-6066.CIR-18-0136
11. Roh W, Chen PL, Reuben A, et al. Integrated molecular analysis of tumor biopsies on sequential CTLA-4 and PD-1 blockade reveals markers of response and resistance. *Sci Transl Med*. 03 2017;9(379)doi:10.1126/scitranslmed.aah3560
12. Amaria RN, Reddy SM, Tawbi HA, et al. Neoadjuvant immune checkpoint blockade in high-risk resectable melanoma. *Nat Med*. 11 2018;24(11):1649-1654. doi:10.1038/s41591-018-0197-1
13. Yost KE, Satpathy AT, Wells DK, et al. Clonal replacement of tumor-specific T cells following PD-1 blockade. *Nat Med*. 08 2019;25(8):1251-1259. doi:10.1038/s41591-019-0522-3
14. Gerlinger M, Rowan AJ, Horswell S, et al. Intratumor heterogeneity and branched evolution revealed by multiregion sequencing. *N Engl J Med*. Mar 8 2012;366(10):883-92. doi:10.1056/NEJMoa1113205
15. Rosenthal R, Cadieux EL, Salgado R, et al. Neoantigen-directed immune escape in lung cancer evolution. *Nature*. Mar 2019;567(7749):479-485. doi:10.1038/s41586-019-1032-7

16. McGranahan N, Furness AJ, Rosenthal R, et al. Clonal neoantigens elicit T cell immunoreactivity and sensitivity to immune checkpoint blockade. *Science*. Mar 25 2016;351(6280):1463-9. doi:10.1126/science.aaf1490
17. Miao D, Margolis CA, Vokes NI, et al. Genomic correlates of response to immune checkpoint blockade in microsatellite-stable solid tumors. *Nat Genet*. Sep 2018;50(9):1271-1281. doi:10.1038/s41588-018-0200-2
18. Joshi K, de Massy MR, Ismail M, et al. Spatial heterogeneity of the T cell receptor repertoire reflects the mutational landscape in lung cancer. *Nat Med*. 10 2019;25(10):1549-1559. doi:10.1038/s41591-019-0592-2
19. Gerlinger M, Quezada SA, Peggs KS, et al. Ultra-deep T cell receptor sequencing reveals the complexity and intratumour heterogeneity of T cell clones in renal cell carcinomas. *J Pathol*. Dec 2013;231(4):424-32. doi:10.1002/path.4284
20. Chen Z, Zhang C, Pan Y, et al. T cell receptor  $\beta$ -chain repertoire analysis reveals intratumour heterogeneity of tumour-infiltrating lymphocytes in oesophageal squamous cell carcinoma. *J Pathol*. Aug 2016;239(4):450-8. doi:10.1002/path.4742
21. Reuben A, Gittelman RM, Gao J, et al. TCR Repertoire Intratumor Heterogeneity in Localized Lung Adenocarcinomas: an Association with Predicted Neoantigen Heterogeneity and Postsurgical Recurrence. *Cancer Discov*. Jul 2017;doi:10.1158/2159-8290.CD-17-0256
22. Reuben A, Spencer CN, Prieto PA, et al. Genomic and immune heterogeneity are associated with differential responses to therapy in melanoma. *NPJ Genom Med*. 2017;2doi:10.1038/s41525-017-0013-8
23. Feng L, Qian H, Yu X, et al. Heterogeneity of tumor-infiltrating lymphocytes ascribed to local immune status rather than neoantigens by multi-omics analysis of glioblastoma multiforme. *Sci Rep*. Jul 2017;7(1):6968. doi:10.1038/s41598-017-05538-z
24. Wang T, Wang C, Wu J, et al. The Different T-cell Receptor Repertoires in Breast Cancer Tumors, Draining Lymph Nodes, and Adjacent Tissues. *Cancer Immunol Res*. Feb 2017;5(2):148-156. doi:10.1158/2326-6066.CIR-16-0107
25. Dash P, Fiore-Gartland AJ, Hertz T, et al. Quantifiable predictive features define epitope-specific T cell receptor repertoires. *Nature*. 07 2017;547(7661):89-93. doi:10.1038/nature22383
26. Danilova L, Anagnostou V, Caushi JX, et al. The Mutation-Associated Neoantigen Functional Expansion of Specific T Cells (MANAFEST) Assay: A Sensitive Platform for Monitoring Antitumor Immunity. *Cancer Immunol Res*. 08 2018;6(8):888-899. doi:10.1158/2326-6066.CIR-18-0129
27. Chen H, Ye F, Guo G. Revolutionizing immunology with single-cell RNA sequencing. *Cell Mol Immunol*. 03 2019;16(3):242-249. doi:10.1038/s41423-019-0214-4
28. Fairfax BP, Taylor CA, Watson RA, et al. Peripheral CD8. *Nat Med*. 02 2020;26(2):193-199. doi:10.1038/s41591-019-0734-6
29. Zhang Z, Xiong D, Wang X, Liu H, Wang T. Mapping the functional landscape of T cell receptor repertoires by single-T cell transcriptomics. *Nat Methods*. 01 2021;18(1):92-99. doi:10.1038/s41592-020-01020-3
30. Li T, Zhao L, Yang Y, et al. T Cells Expanded from PD-1. *Cancer Res*. Jan 2021;doi:10.1158/0008-5472.CAN-20-2300
31. Glanville J, Huang H, Nau A, et al. Identifying specificity groups in the T cell receptor repertoire. *Nature*. 07 2017;547(7661):94-98. doi:10.1038/nature22976
32. Ostmeier J, Christley S, Toby IT, Cowell LG. Biophysicochemical Motifs in T-cell Receptor Sequences Distinguish Repertoires from Tumor-Infiltrating Lymphocyte and Adjacent Healthy Tissue. *Cancer Res*. 04 2019;79(7):1671-1680. doi:10.1158/0008-5472.CAN-18-2292
33. Leem J, de Oliveira SHP, Krawczyk K, Deane CM. STCRDab: the structural T-cell receptor database. *Nucleic Acids Res*. 01 2018;46(D1):D406-D412. doi:10.1093/nar/gkx971
34. Singh NK, Abualrous ET, Ayres CM, et al. Geometrical characterization of T cell receptor binding modes reveals class-specific binding to maximize access to antigen. *Proteins*. 03 2020;88(3):503-513. doi:10.1002/prot.25829

35. Thomas N, Best K, Cinelli M, et al. Tracking global changes induced in the CD4 T-cell receptor repertoire by immunization with a complex antigen using short stretches of CDR3 protein sequence. *Bioinformatics*. Nov 2014;30(22):3181-8. doi:10.1093/bioinformatics/btu523
36. Pogorelyy MV, Minervina AA, Shugay M, et al. Detecting T cell receptors involved in immune responses from single repertoire snapshots. *PLoS Biol*. 06 2019;17(6):e3000314. doi:10.1371/journal.pbio.3000314
37. Marcou Q, Mora T, Walczak AM. High-throughput immune repertoire analysis with IGoR. *Nat Commun*. 02 2018;9(1):561. doi:10.1038/s41467-018-02832-w
38. Huang H, Wang C, Rubelt F, Scriba TJ, Davis MM. Analyzing the Mycobacterium tuberculosis immune response by T-cell receptor clustering with GLIPH2 and genome-wide antigen screening. *Nat Biotechnol*. 10 2020;38(10):1194-1202. doi:10.1038/s41587-020-0505-4
39. Chiou SH, Tseng D, Reuben A, et al. Global analysis of shared T cell specificities in human non-small cell lung cancer enables HLA inference and antigen discovery. *Immunity*. Mar 2021;54(3):586-602.e8. doi:10.1016/j.immuni.2021.02.014
40. Sidhom JW, Larman HB, Pardoll DM, Baras AS. DeepTCR is a deep learning framework for revealing sequence concepts within T-cell repertoires. *Nat Commun*. 03 2021;12(1):1605. doi:10.1038/s41467-021-21879-w
41. Springer I, Besser H, Tickotsky-Moskovitz N, Dvorkin S, Louzoun Y. Prediction of Specific TCR-Peptide Binding From Large Dictionaries of TCR-Peptide Pairs. *Front Immunol*. 2020;11:1803. doi:10.3389/fimmu.2020.01803
42. Moris P, De Pauw J, Postovskaya A, et al. Current challenges for unseen-epitope TCR interaction prediction and a new perspective derived from image classification. *Brief Bioinform*. Dec 2020;doi:10.1093/bib/bbaa318
43. Tong Y, Wang J, Zheng T, et al. SETE: Sequence-based Ensemble learning approach for TCR Epitope binding prediction. *Comput Biol Chem*. Jun 2020;87:107281. doi:10.1016/j.compbiolchem.2020.107281
44. Borrman T, Pierce BG, Vreven T, Baker BM, Weng Z. High-throughput modeling and scoring of TCR-pMHC complexes to predict cross-reactive peptides. *Bioinformatics*. Dec 2020;doi:10.1093/bioinformatics/btaa1050
45. Milighetti M, Shawe-Taylor J, Chain B. Predicting T cell receptor antigen specificity from structural features derived from homology models of receptor-peptide-major histocompatibility complexes. bioRxiv 2021.05.19.4448432021.
46. Lanzarotti E, Marcatili P, Nielsen M. Identification of the cognate peptide-MHC target of T cell receptors using molecular modeling and force field scoring. *Mol Immunol*. 02 2018;94:91-97. doi:10.1016/j.molimm.2017.12.019

## Figure Legends

**Figure 1. T cell clonality and diversity.** T cells represented by coloured spheres. The colour of each T cell represents its clonotype. In a tumour sample where T cell clonality is low any diversity is high, many T cell clones with different specificities are present and none are very expanded (left). In contrast, in a sample with increased T cell clonality and low T cell diversity, only a few T cell clones are detected and are highly expanded (right).

## Figure 2. Applications of TCR repertoire analysis in cancer immunology.

**A) TCR response to immune checkpoint blockade.** T cells represented by coloured spheres. The colour of each T cell represents its clonotype. Anti-CTLA-4 treatment is associated with increased TCR repertoire diversity (left) whereas anti-PD-1/PD-L1 treatment is associated with oligoclonal T cell expansion.

**B) TCR repertoire heterogeneity reflects tumoural genomic heterogeneity.** T cells represented as in A. Tumour cells are star shaped and different colours represent different tumour subclones based on their genomic landscape.

**C) Assessment of TCR antigen specificity.** By combining functional assays with TCR sequencing, TILs can be annotated with their antigen specificity to provide further insights into the anti-tumoural immune response and provide candidates for TCR therapy. TILs are known to include neoantigen-reactive T cells (NARTs), which respond to mutant proteins unique to the tumour, T cells specific for tumour-associated antigens (TAAs) and bystander T cells, which do not recognise tumour.

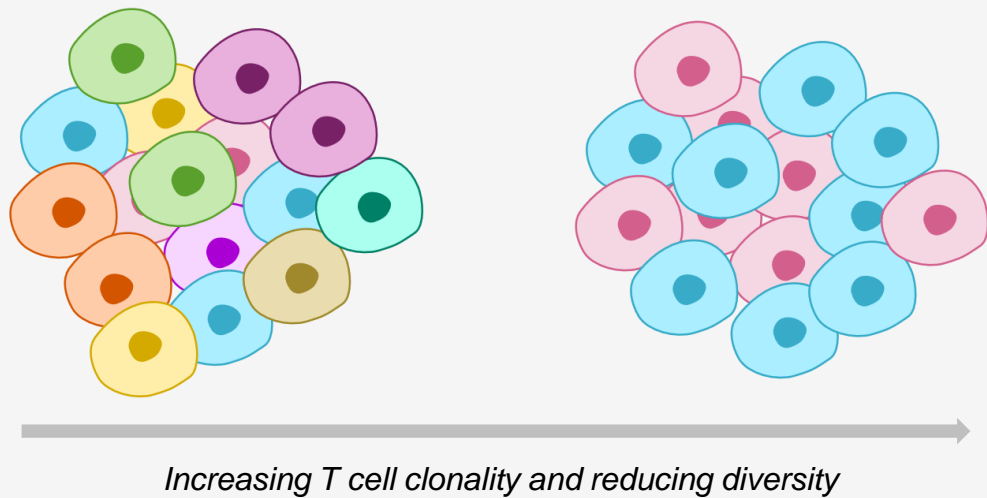
## Figure 3. Computational methods for TCR-antigen pairing.

**A) Clustering algorithms.** Clustering algorithms find TCRs within a repertoire that are similar in sequence and therefore may share antigenic specificity. A distance metric is calculated between each pair of TCRs based on their sequence and they are assigned to the same cluster if the distance is below a certain threshold.

**B) TCR specificity clusters.** An extension to the clustering algorithms is to assign specificity to each cluster, i.e., a peptide antigen that the TCRs in the cluster are known to recognise. New TCRs with unknown specificity can then be assigned to a specificity cluster depending on the distance calculated from the members of each cluster.

**C) Prediction of binding probability.** A classifier is trained on a set of known TCR-antigen pairs, as well as a set of known non-binders. For a new TCR-antigen pair, a classifier will output a probability that the TCR will bind the specific peptide.

**“Declarations of interest: none.”**

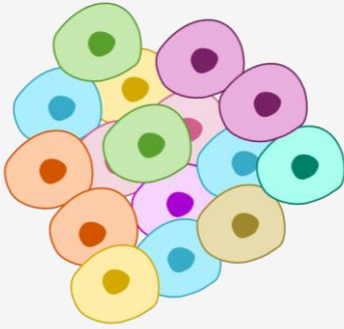


**Figure 1. T cell clonality and diversity.** T cells represented by coloured spheres. The colour of each T cell represents its clonotype. In a tumour sample where T cell clonality is low any diversity is high, many T cell clones with different specificities are present and none are very expanded (left). In contrast, in a sample with increased T cell clonality and low T cell diversity, only a few T cell clones are detected and are highly expanded (right).

## A) Response to immune checkpoint blockade

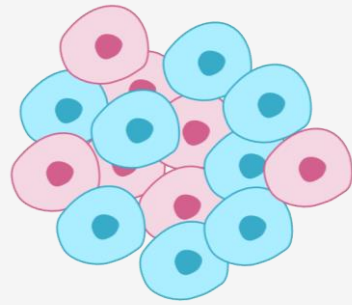
Anti-CTLA-4

*Increased repertoire diversity*

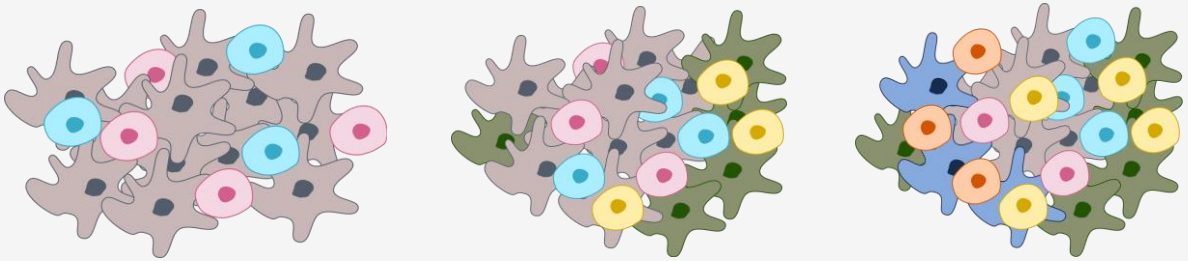


Anti-PD-L1/PD-1

*Oligoclonal T cell expansion*

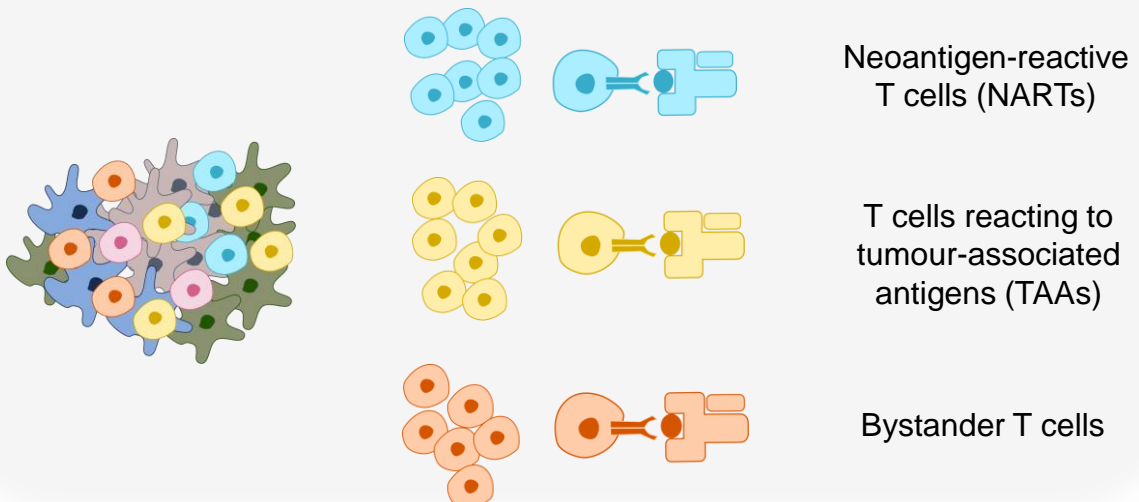


## B) Intra-tumoural heterogeneity of the TCR repertoire reflects genomic heterogeneity



*Increasing tumour genomic heterogeneity and increasing TCR repertoire heterogeneity*

## C) Assessment of TCR antigenic specificity





## Figure 2. Applications of TCR repertoire analysis in cancer immunology.

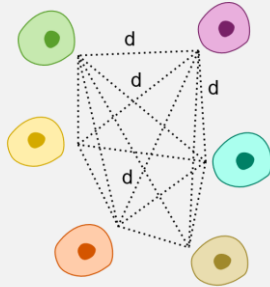
**A) TCR response to immune checkpoint blockade.** T cells represented by coloured spheres. The colour of each T cell represents its clonotype. Anti-CTLA-4 treatment is associated with increased TCR repertoire diversity (left) whereas anti-PD-1/PD-L1 treatment is associated with oligoclonal T cell expansion.

**B) TCR repertoire heterogeneity reflects tumoural genomic heterogeneity.** T cells represented as in A. Tumour cells are star shaped and different colours represent different tumour subclones based on their genomic landscape.

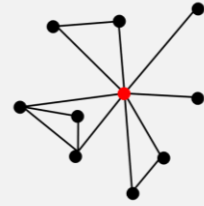
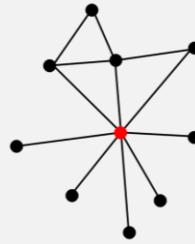
**C) Assessment of TCR antigen specificity.** By combining functional assays with TCR sequencing, TILs can be annotated with their antigen specificity to provide further insights into the anti-tumoural immune response and provide candidates for TCR therapy. TILs are known to include neoantigen-reactive T cells (NARTs), which respond to mutant proteins unique to the tumour, T cells specific for tumour-associated antigens (TAAs) and bystander T cells, which do not recognise tumour.

## A) Clustering algorithms

Identification of TCRs possibly responding to the same (unknown) antigen from bulk sequencing by calculation of sequence similarity



Clustering



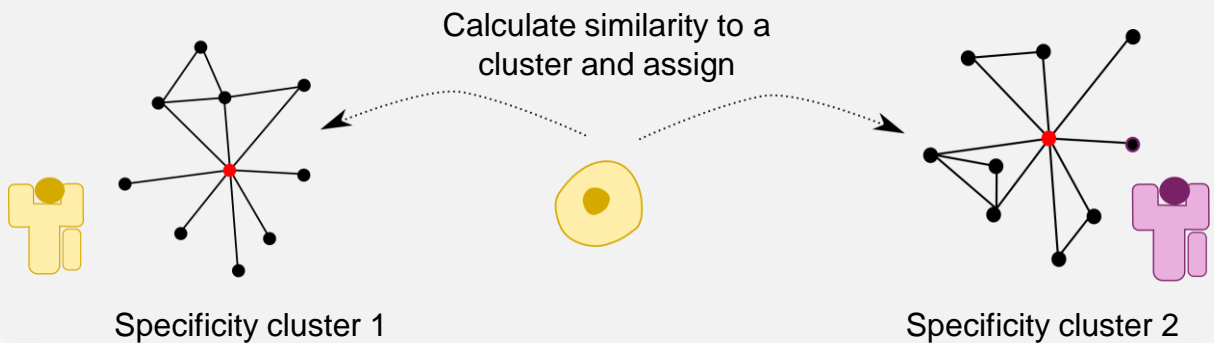
Calculate all pairwise distances

Specificity cluster 1

Specificity cluster 2

## B) TCR specificity clusters

TCR specificity groups are created based on a similarity metrics. A TCR with unknown specificity can be assigned to an existing group because of its similarity to the members of the group (triplet kernel, TCRdist, GLIPH, ...)



## C) Prediction of probability of binding

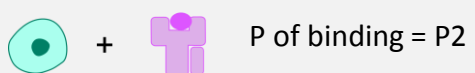
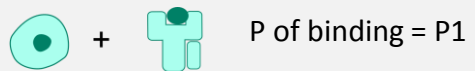
TCR-pMHC pairs with known binding are used to train a predictive model. For a TCR with unknown specificity, probability of binding a specific pMHC is calculated (ERGO, ImRex, ...)

Set of known binding pairs



Machine Learning

Prediction for unknown pairs



### **Figure 3. Computational methods for TCR-antigen pairing.**

**A) Clustering algorithms.** Clustering algorithms find TCRs within a repertoire that are similar in sequence and therefore may share antigenic specificity. A distance metric is calculated between each pair of TCRs based on their sequence and they are assigned to the same cluster if the distance is below a certain threshold.

**B) TCR specificity clusters.** An extension to the clustering algorithms is to assign specificity to each cluster, i.e., a peptide antigen that the TCRs in the cluster are known to recognise. New TCRs with unknown specificity can then be assigned to a specificity cluster depending on the distance calculated from the members of each cluster.

**C) Prediction of binding probability.** A classifier is trained on a set of known TCR-antigen pairs, as well as a set of known non-binders. For a new TCR-antigen pair, a classifier will output a probability that the TCR will bind the specific peptide.

## Highlights

- TCR repertoire analysis as a tool to advance cancer immunotherapy
- Computational methods are used to understand TCR repertoire characteristics
- TCR sequence and structure can be used for specificity prediction

13<sup>th</sup> July 2021

Dear Hans

Thank you for reviewing our MS and for the helpful comments. We have edited the MS to address these points as follows:

1. On page 3 the authors talk about the observation that CTLA-4 blockage leads to increase TCR diversity. This requires further explanation; is it simply a consequence of local T cell recruitment (as suggested); does 'new recruitment' also apply to blood; are new TCR generated in the thymus? This is a very interesting point. Remarkably, the systemic effects of anti-CTLA4 on the overall repertoire seem to be remarkably limited, and the effects on intra-tumour heterogeneity more marked. Perhaps the concentrations achieved in vivo are such that the effect on the immune system as a whole (blood, thymus etc.) are relatively limited, and the effects are more marked where Tregs are actively limiting an immune response such as in a tumour. We have added a couple of sentences to explain this point on page 3
2. In the section 'Repertoire analysis in assigning TCR antigenic specificity' to authors should clarify whether cited papers use CDR3 motifs in one TCR chain, or combine CDR3 motifs in the alpha and beta chain. The authors should also clarify whether CDR1 and CDR2 are included in any of the specificity prediction models.

Although the prediction tools do not usually explicitly make use of the other CDRs, the structure-based ones implicitly make use of the whole TCR sequence. We have added a section to address this point at the top of page 7.

3. The left panel in figure 2B suggests that low mutation rates drive clonality of tumour infiltrating T cells. Is this the case? No, we do did not wish to imply this; only that increased mutational heterogeneity leads to increased TCR heterogeneity. We have edited the figure somewhat, and added an explanatory sentence in the text (top of page 5).

I hope the revised manuscript is now acceptable for publication.

Many thanks



With best wishes

Benny Chain  
Professor of Computational Immunology

UCL Division of Infection and Immunity  
Cruciform Building, Gower St., London WC16BT  
Tel: +44 (0) 2031082129  
b.chain@ucl.ac.uk  
[www.ucl.ac.uk/infection-immunity](http://www.ucl.ac.uk/infection-immunity)



[Click here to access/download](#)

**Annotated references**

[annotated\\_references\\_MM2\\_KJ3\\_BC\\_final.docx](#)

