



Article

# Genetic Bias, Diversity Indices, Physicochemical Properties and CDR3 Motifs Divide Auto-Reactive from Allo-Reactive T-Cell Repertoires

Oscar L. Haigh <sup>1</sup>, Emma J. Grant <sup>2</sup>, Thi H. O. Nguyen <sup>3</sup> , Katherine Kedzierska <sup>3</sup> , Matt A. Field <sup>4,5,6,7,†</sup> and John J. Miles <sup>4,5,6,8,\*</sup>

- <sup>1</sup> Center for Immunology of Viral, Auto-Immune, Hematological and Bacterial Diseases (IMVA-HB), Commissariat à l'Énergie Atomique et aux Énergies renouvelables, Université Paris-Saclay, INSERM U1184, Fontenay-aux-Roses, France; oscar.haigh@cea.fr
  - <sup>2</sup> Department of Biochemistry and Genetics, La Trobe Institute for Molecular Science, La Trobe University, Melbourne, VIC 3086, Australia; E.Grant@latrobe.edu.au
  - <sup>3</sup> Department of Microbiology and Immunology, at the Peter Doherty Institute for Infection and Immunity, University of Melbourne, Parkville, VIC 3010, Australia; tho.nguyen@unimelb.edu.au (T.H.O.N.); kkedz@unimelb.edu.au (K.K.)
  - <sup>4</sup> The Australian Institute of Tropical Health and Medicine (AITHM), James Cook University, Cairns, QLD 4811, Australia; matt.field@jcu.edu.au
  - <sup>5</sup> Centre for Molecular Therapeutics, James Cook University, Cairns, QLD 4811, Australia
  - <sup>6</sup> Centre for Tropical Bioinformatics and Molecular Biology, James Cook University, Cairns, QLD 4811, Australia
  - <sup>7</sup> John Curtin School of Medical Research, Australian National University, Canberra, ACT 2601, Australia
  - <sup>8</sup> QIMR Berghofer Medical Research Institute, Brisbane, QLD 4006, Australia
- \* Correspondence: john.miles@jcu.edu.au; Tel.: +617-4232-1284  
† These authors contributed equally to this work.



**Citation:** Haigh, O.L.; Grant, E.J.; Nguyen, T.H.O.; Kedzierska, K.; Field, M.A.; Miles, J.J. Genetic Bias, Diversity Indices, Physicochemical Properties and CDR3 Motifs Divide Auto-Reactive from Allo-Reactive T-Cell Repertoires. *Int. J. Mol. Sci.* **2021**, *22*, 1625. <https://doi.org/10.3390/ijms22041625>

Academic Editor: Jeffrey L. Platt  
Received: 4 December 2020  
Accepted: 31 January 2021  
Published: 5 February 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Abstract:** The distinct properties of allo-reactive T-cell repertoires are not well understood. To investigate whether auto-reactive and allo-reactive T-cell repertoires encoded distinct properties, we used dextramer enumeration, enrichment, single-cell T-cell receptor (TCR) sequencing and multiparameter analysis. We found auto-reactive and allo-reactive T-cells differed in mean ex vivo frequency which was antigen dependent. Allo-reactive T-cells showed clear differences in TCR architecture, with enriched usage of specific T-cell receptor variable (*TRBJ*) genes and broader use of T-cell receptor variable joining (*TRBJ*) genes. Auto-reactive T-cell repertoires exhibited complementary determining regions three (CDR3) lengths using a Gaussian distribution whereas allo-reactive T-cell repertoires exhibited distorted patterns in CDR3 length. CDR3 loops from allo-reactive T-cells showed distinct physical-chemical properties, tending to encode loops that were more acidic in charge. Allo-reactive T-cell repertoires differed in diversity metrics, tending to show increased overall diversity and increased homogeneity between repertoires. Motif analysis of CDR3 loops showed allo-reactive T-cell repertoires differed in motif preference which included broader motif use. Collectively, these data conclude that allo-reactive T-cell repertoires are indeed different to auto-reactive repertoires and provide tangible metrics for further investigations and validation. Given that the antigens studied here are overexpressed on multiple cancers and that allo-reactive TCRs often show increased ligand affinity, this new TCR bank also has translational potential for adoptive cell therapy, soluble TCR-based therapy and rational TCR design.

**Keywords:** T-cell; T-cell receptor; T-cell repertoire; T-cell alloreactivity

## 1. Introduction

Many questions remain on the distinct properties of allo-reactive T-cells [1]. Despite the fact that allo-reactive T-cells are an established clinical correlate of graft rejection, knowledge of the precise molecular characteristics of this T-cell repertoire are largely

unknown. Initial T-cell receptor (TCR) repertoire studies have shown TCR repertoire structure correlates with rejection [2,3], cancer protection post-transplant [4], pathogen infection post-transplant [3], and high TCR avidity [5].

Three models of alloreactivity have been proposed [6]. First, is that allo-reactive TCRs recognise foreign human leukocyte antigen (HLA) molecules directly, with the peptide (p) being irrelevant, termed the “promiscuous model”. Second, is that allo-reactive TCRs rearrange their CDR3 loops to recognise promiscuous shapes, termed the “degenerate model”. Third, is that allo-reactive TCRs bind both peptide and HLA in conventional form, making small changes in the contact interface that dramatically modify binding affinity, termed the “polyspecific model”. Massing functional, biophysical, and structural evidence supports the polyspecific model with allo-reactive TCRs binding both HLA and peptide in a conventional manner, often through molecular mimicry binding “hotspot” regions [7–10].

Solid-organ transplant (SOT) and hematopoietic stem-cell transplantation (HSCT) rejection can be controlled using T-cell suppressing drugs such as cyclosporine, corticosteroids, rapamycin or infusions with T-cell suppressing monoclonal antibodies. However, these life-long T-cell suppressing drug regimens come with increased risk of infection, cardiovascular disease and cancer [11]. Acute graft rejection may occur between one-week and three-weeks post-transplant. Chronic rejection can occur over many years. It is thought that chronic rejection can be triggered by innocuous infections, which activate allo-reactive memory T-cells that damage host tissue [12]. Many SOT rejection episodes involve T cell-mediated damage, but these episodes do not always destroy grafts [13,14].

T-cell mediated HSCT rejection comprises three modalities. First, T-cells can recognise host pHLA directly, termed the “direct alloresponse”. Second, T-cell-derived peptide fragments presented by donor HLA can be recognised, termed the “indirect alloresponse”. Third, T-cells can recognise pMHC molecules that have been transferred from donor cells to recipient cells, termed the “semidirect alloresponse” [14].

Currently, structural, biophysical, mutational, and functional analysis of auto-reactive and allo-reactive TCRs and pHLAs agrees with a conventional top-down docking mode, where both the germline loops of the TCR and the somatic complementary determining regions three (CDR3) loops contact both the peptide and HLA. This docking occurs in a roughly diagonal manner with binding affinities between  $K_D = 0.3 \mu\text{M}$  and above 1 mM [9,10,15,16]. However, this conventional docking mode can sometimes vary [17]. The mouse major histocompatibility complex (MHC-I) restricted 2C TCR binds its auto-reactive ligand at  $K_D = 83\text{--}85 \mu\text{M}$ , approximately 40-fold weaker its allo-reactive ligands at  $K_D = 2\text{--}3 \mu\text{M}$  [10,18,19]. In humans, the antiviral LC13 TCR binds its auto-reactive ligand at  $K_D = 10 \mu\text{M}$  [16], approximately 10-fold weaker than its allo-reactive ligand at  $K_D = 1 \mu\text{M}$  [17]. The structures of these complexes show they bind both HLA and peptide with conventional footprints [17–23]. Of additional interest is the auto-reactive MHC-II-restricted mouse TCR YAe62 can also bind a MHC-I ligand [24]. This TCR transcends MHC-restriction dogma. The same phenomenon has been shown in humans, with a virus-specific HLA-I-restricted TCR being allo-reactive with an HLA-II-restricted ligand [25]. TCRs are also known to recognise different antigens presented by different HLA-I alleles [26,27]. In general, the polyspecific model appears to best conform with alloreactivity [22], with the peptide being the focus of affinity variation [28] through increased atomic contacts [22]. However, TCR/pMHC structural correlations or multimer correlations have not yet been explicitly linked with graft acceptance or rejection in humans.

In this study, we aimed to investigate two gaps in knowledge. First, was to determine whether antigen-specific precursors differed in frequency between auto-reactive and allo-reactive individuals. Second, was to determine if T-cell repertoires could be separated into auto-reactive and allo-reactive groups when quantitating TCR gene preference, TCR gene pairings, CDR3 length, CDR3 physiochemical-properties, CDR3 diversity metrics and CDR3 motifs. Here, we used HLA-0201 dextramers bearing four different immunogenic tumour associated antigens (TAAs). The first was IMNDMPIYM (herein referred to as IMN) derived from the transmembrane receptor EphA2 and overexpressed in prostate, breast,

colon, gastric cancers, and metastatic melanoma [29–31]. Second, was FLCMKALLL (herein referred to as FLC) [32] derived from the android receptor ligand-binding domain protein and overexpressed in prostate cancer, breast cancers and hepatocellular carcinoma [33–35]. Third, was KVAELVHFL (herein referred to as KVA) derived from the MAGE-A3 protein and overexpressed in prostate, breast, and esophageal cancers [36–38]. Fourth, was YLEPGPVTV (herein referred to as YLE) derived from the gp100 protein, which is found in premelanosomal pigmented cells and the retina and overexpressed in numerous cancers, including prostate cancer [38]. In this study, HLA-0201<sup>+</sup> donors are referred to as auto-reactive and HLA-0201<sup>−</sup> donors are referred to as allo-reactive.

## 2. Results

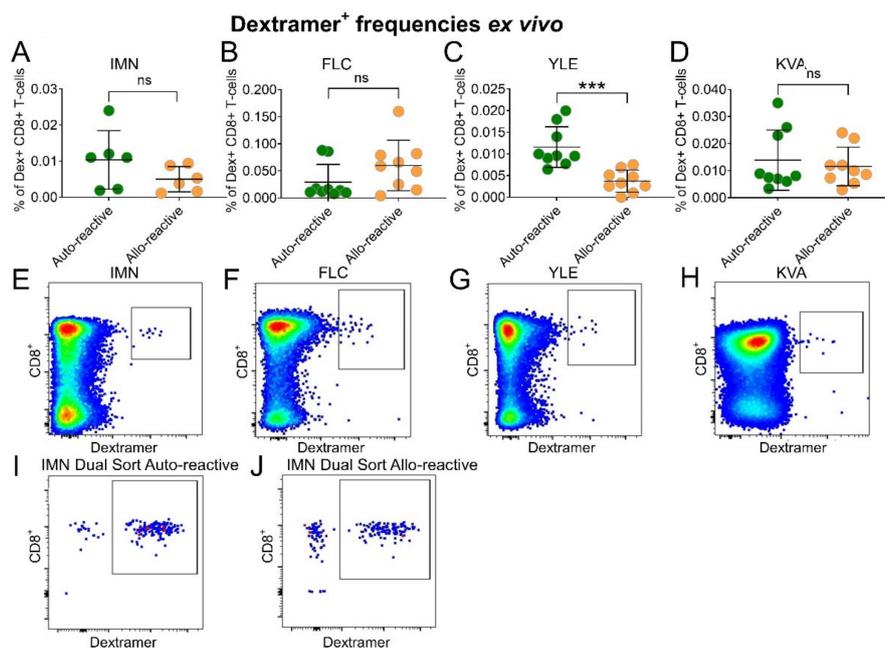
### 2.1. Antigen-Specific T-Cells Show Different Frequencies between Auto-Reactive and Allo-Reactive Settings

Previous experience enumerating and sorting extremely low-frequency T-cell populations in cord and adult blood [39] has shown us success requires; (i) the use of a bright fluorochrome on the dextramer [40]; (ii) the use of freshly made dextramers which increases brightness 13-to-50-fold, compared with other multimers [39,41]); (iii) the use of the protein kinase inhibitor dasatinib which prevents TCR internalisation, activation-induced cell death [42] and increases multimer brightness up to 50-fold [40]; (iv) cell dumping using LIVE/DEAD, CD14, CD16 and CD19 [40] and; (v) a decreased sorter flow rate which minimises the abort rate of dextramer+ cells. An important variable in this protocol is the functional avidity of the T-cell population under investigation, which derives from the physiology and activation status of the cells.

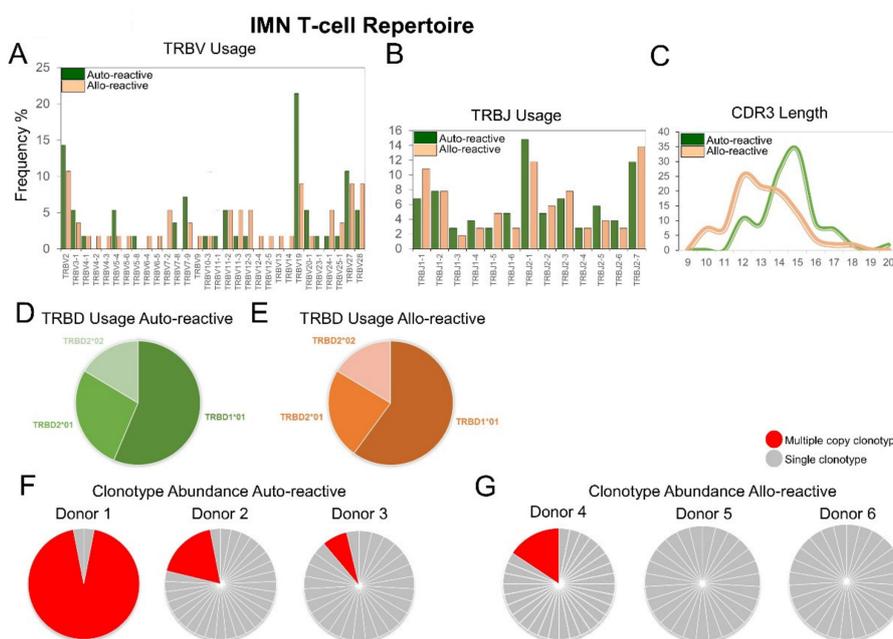
Using flow cytometry, we found that IMN- and FLC-specific T-cells had variable mean frequencies across auto-reactive and allo-reactive populations (Figure 1A,B), YLE-specific T-cells had a significantly higher mean frequency ( $p \leq 0.0002$ ) in auto-reactive populations than allo-reactive populations (Figure 1C). KVA-specific T-cells exhibited a higher mean frequency when comparing auto-reactive and allo-reactive populations (Figure 1D). Collectively, the frequency hierarchy was FLC > IMN > KVA > YLE, ranging from a mean of 0.004–0.06% of CD8<sup>+</sup> T-cells. Examples of dextramer plots from ex vivo samples are shown (Figure 1E–H), revealing the low percentages of these cells in the blood. Examples of samples that have been sorted twice for purity (and further single-cell sorting) are also shown (Figure 1I,J), highlighting that antigen-specific T-cells can be reliably isolated and enriched from both auto-reactive and allo-reactive settings.

### 2.2. The IMN Allo-Reactive T-Cell Repertoire Exhibits Expanded TRBV Gene Usage and Unconventional CDR3 Loop Length

We next performed single-cell sorting and TCR analysis on IMN- and FLC-specific T-cells given that these dextramers had the highest mean fluorescence intensity (MFI) of the dextramer set and had the highest frequencies in the blood. These features help confirm precise sorting and provided confidence in downstream analysis. IMN-specific TCR sequences were then pooled and normalised by clonotype count and appearance. Differences in TRBV usage was highly significant between cohorts showing 77% of total variation by two-way ANOVA ( $p \leq 0.0001$ ) (Figure 2A). Here, 27-of-30 (90%) TRBV genes were deployed in allo-reactive repertoires versus 19-of-30 (63%) in auto-reactive repertoires. The total variation in TRBJ usage was also highly significant between cohorts at 92% by two-way ANOVA ( $p \leq 0.0001$ ) (Figure 2B). There was also a distinct difference in CDR3 length between repertoires with the allo-reactive repertoire not exhibiting a conventional Gaussian distribution in CDR3 length, which typically shows a peak at 15 residues in length. The CDR3 length variation between cohorts was 76% and significant by two-way ANOVA ( $p \leq 0.025$ ) (Figure 2C). Interestingly, allo-reactive TCRs tended to encode shorter CDR3 loops. Here, in allo-reactive CDR3 loops, 61% of the repertoire was under 13 residues in length versus 18% in the auto-reactive cohort. There were no significant differences in TRBD usage between auto-reactive (Figure 2D) and allo-reactive (Figure 2E) TCRs, with the TRBD1\*01 gene predominating in both cohorts.



**Figure 1.** Different frequencies among dextramer populations in the blood for auto-reactive and allo-reactive T-cells. Dextramers include (A) IMN, (B) FLC, (C) YLE and (D) KVA. Frequency of CD8+ cells is shown on the y-axis, and dextramer frequencies from auto-reactive (green) and allo-reactive (orange) donors on the x-axis. Comparisons of IMN, FLC and KVA cohorts were not significant. Frequencies from the YLE-specific cohorts were significantly different, with the auto-reactive donors dominating. Examples of IMN (E), FLC (F), YLE (G) and KVA (H) dextramer staining ex vivo. Examples of dual (sequential) dextramer sorting from an (I) auto-reactive repertoire and (J) allo-reactive repertoire using IMN. ns, no significant; \*\*\*  $p \leq 0.001$ .

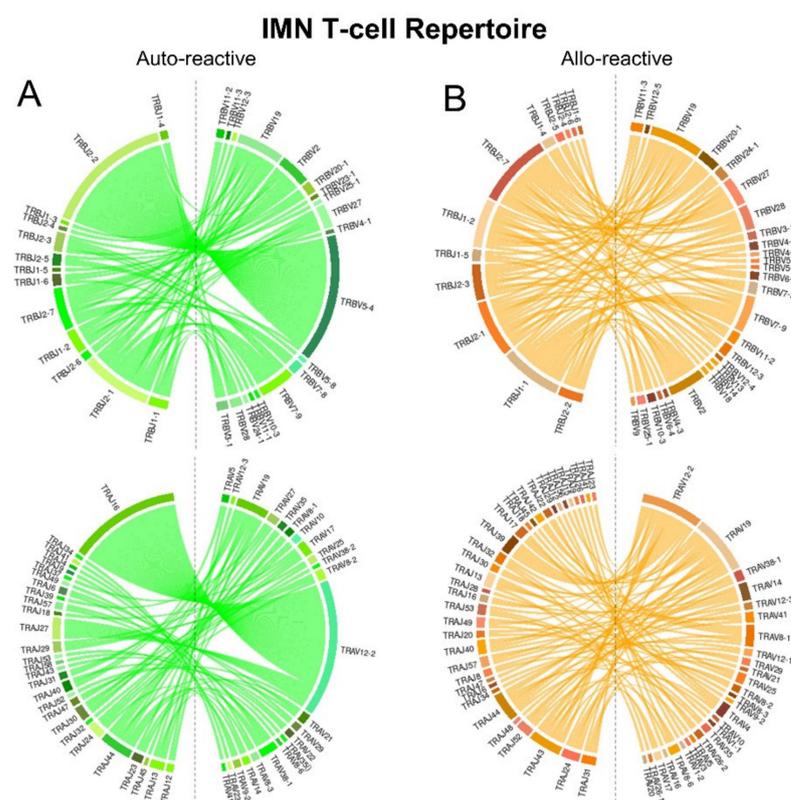


**Figure 2.** Different IMN-specific T-cell repertoires between auto-reactive donors and allo-reactive donors. Bar charts showing percentage use of (A) TRBV and (B) TRBJ preference from auto-reactive (green) versus allo-reactive (orange) donors. (C) Line chart comparing percentage use of CDR3 length. (D,E) Pie charts comparing percentage use of the TRBD genes. Pie charts showing clonotype frequencies in (F) three auto-reactive donors and (G) three allo-reactive donors. Each slice represents a unique clonotype from the repertoire of a single donor. Unique clonotypes with multiple copies are shown in red and unique clonotypes with a single copy are shown in grey. Percentages were determined by summing the total clonotypes found within a single donor.

We then compared total clonotype frequencies in three auto-reactive donors (Figure 2F) and three allo-reactive donors (Figure 2G) by total abundance sampling. Here, three-of-three auto-reactive donors exhibited multiple copies of an individual clonotype against IMN, with frequencies up to 93% for a single clonotype in donor 1 (Figure 2F). Within the allo-reactive donors, one-of-three individuals exhibited multiple copies of an individual clonotype against IMN, with a frequency of 16% for a single clonotype in donor 4 (Figure 2G).

### 2.3. IMN-Specific T-Cell Repertoires Exhibit Divergent TCR Gene Pairing

IMN-specific T-cell repertoires were next compared for TRBV/*JTRBJ* and TRAV/TRAJ pairing bias using normalised counts by abundance and chord diagram analysis (Figure 3). In the auto-reactive repertoire, the large unique multiple copy clonotype from auto-reactive donor 1 (TRAV12-2/TRBV5-4) was observed along with preferential pairing of TRBV7-9 and *JTRBJ*2-1 for *TRB* genes and TRAV12-2 and TRAJ27 for the *TRA* genes. TCR gene pairings in the allo-reactive repertoire were much more complex. However preferential pairings of TRBV7-9 and TRBV2 with *JTRBJ*2-7 were observed for the *TRB* genes and TRAV12-2 and TRAJ39 for the *TRV* genes. More TRAJ genes were utilized in the allo-reactive repertoire compared with the auto-reactive repertoire.



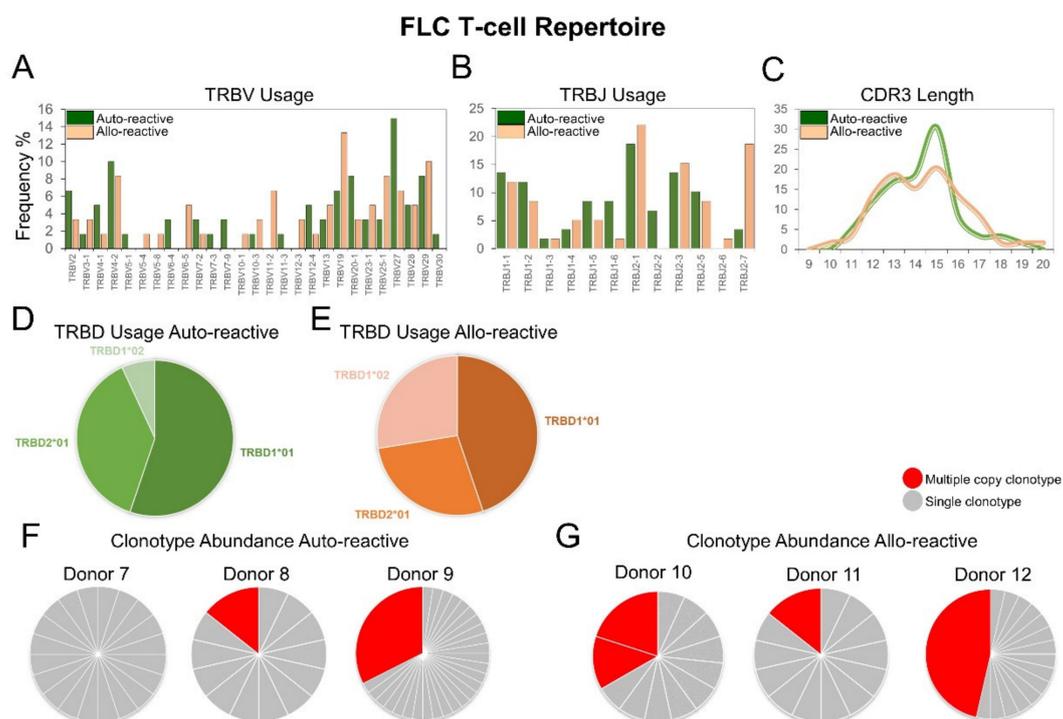
**Figure 3.** Chord diagrams of IMN-specific (A) auto-reactive (green) and (B) allo-reactive (orange) repertoires showing TRBV-*TRBJ* and TRAV-*TRAJ* pairings. The size of the connections represents an increased frequency of appearance.

### 2.4. The FLC-Specific Allo-Reactive T-Cell Repertoire Exhibits Divergent TRBV and TRBJ Gene Usage, Unconventional CDR3 Loop Length and Common Multiple Copy Clonotypes

Like the IMN-specific repertoires, FLC-specific TCR sequences were also pooled and normalised by clonotype count and appearance. Differences in TRBV usage were highly significant showing 73.86% between cohorts by two-way ANOVA ( $p \leq 0.005$ ) (Figure 4A). Differences in *TRBJ* usage were also significant showing 80% variation by two-way ANOVA ( $p \leq 0.016$ ) (Figure 4B). Akin to allo-reactive IMN repertoire, the FLC-specific allo-reactive

repertoire did not exhibit a Gaussian distribution in CDR3 length showing 90% variation compared to the auto-reactive repertoire by two-way ANOVA ( $p \leq 0.0001$ ) (Figure 4C). Allo-reactive CDR3 sequences were shorter than auto-reactive sequences but not to the same degree as seen in the IMN system, with 37% and 34% of total sequences, respectively being under 13 residues in length. There were no significant differences in TRBD usage between auto-reactive (Figure 4D) and allo-reactive (Figure 4E) TCRs, with the TRBD1\*01 gene predominating in both cohorts.

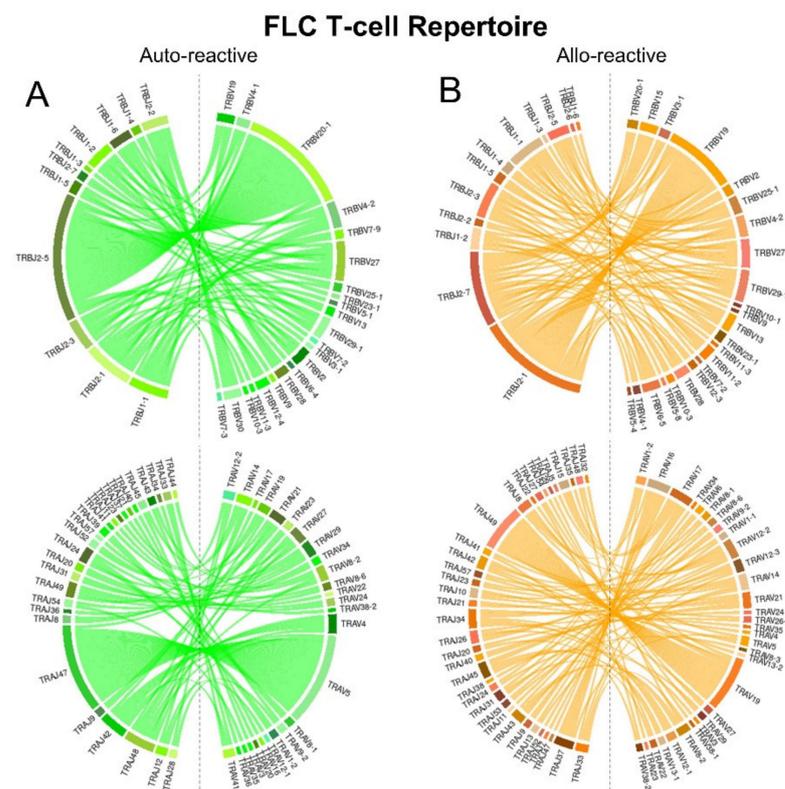
When examining the total clonotype frequencies per donor, two-of-three auto-reactive donors showed unique clonotypes with multiple copies (Figure 4F) versus three-of-three in allo-reactive donors (Figure 4G). Donor 12 showed the largest unique clonotype frequency of 46%. These multiple copy clonotypes in FLC allo-reactive donors are different to the IMN system which were comprised of single unique clonotypes per donor.



**Figure 4.** Different FLC-specific T-cell repertoires between auto-reactive and allo-reactive donors. Bar charts showing percentage use of (A) TRBV and (B) TRBJ preference from auto-reactive (green) versus allo-reactive (orange) donors. (C) Line chart comparing percentage use of CDR3 length. (D,E) Pie charts comparing percentage use of the TRBD genes. Pie charts showing clonotype frequencies in (F) three auto-reactive donors and (G) three allo-reactive donors. Each slice represents a unique clonotype from the repertoire of a single donor. Unique clonotypes with multiple copies are shown in red and unique clonotypes with a single copy are shown in grey. Percentages were determined by summing the total clonotypes found within a single donor.

### 2.5. FLC-Specific T-Cell Repertoires Exhibit Divergent TCR Gene Pairing

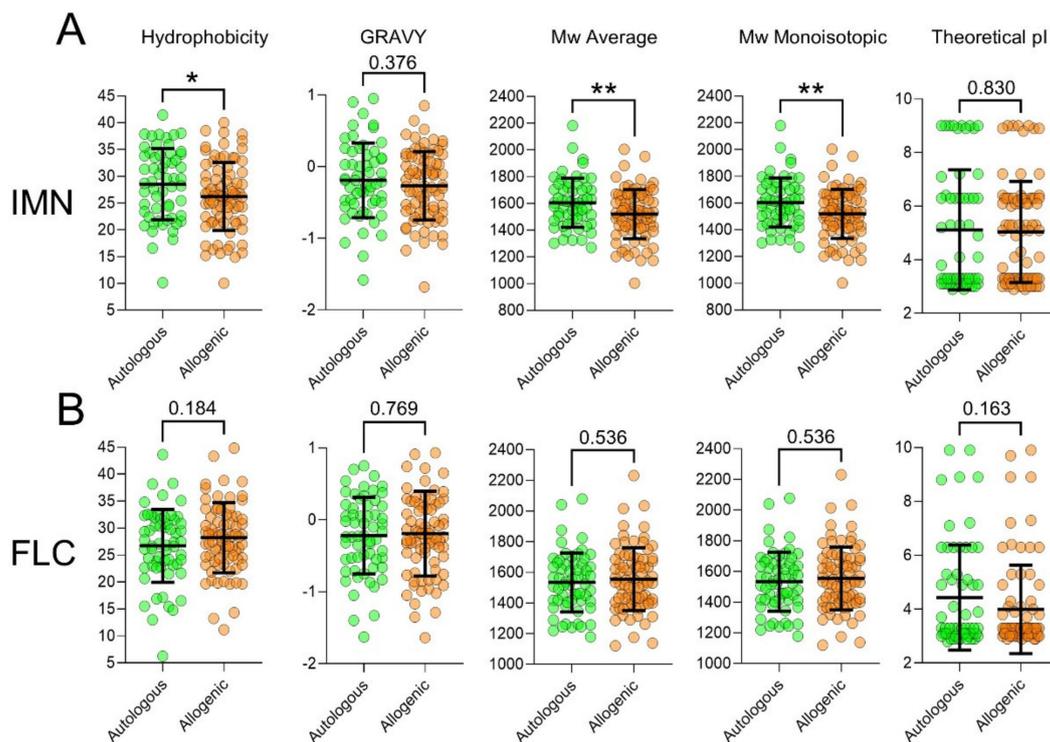
FLC-specific T-cell repertoires were next compared for TRBV/TRBJ and TRAV/TRAJ pairing bias using normalised counts by abundance and chord diagram analysis (Figure 5). In the auto-reactive repertoire, we observed preferential pairing of TRBV20-1 and TRBJ2-5 for TRB genes and TRAV5 and TRAJ47 for the TRA genes. TCR gene pairings in the allo-reactive repertoire were again more complex. However, preferential pairings of TRBV19 and TRBV2-1 were observed for the TRB genes and TRAV19 and TRAJ49 for the TRV genes. Akin to the IMN system, more TRAJ genes were utilized in the allo-reactive repertoire compared with the auto-reactive repertoire.



**Figure 5.** Chord diagrams of FLC-specific (A) auto-reactive (green) and (B) allo-reactive (orange) repertoires showing TRBV-TRBJ and TRAV-TRVJ pairings. The size of the connections represents an increased frequency of appearance.

### 2.6. T-Cells Exhibit Divergent CDR3 Physical-Chemical Properties between Auto-Reactive and Allo-Reactive Repertoires

We next examined the physical-chemical properties of the CDR3 loops in the IMN- and FLC-specific repertoires by pooling the sequences based on appearance. In both systems, all metrics measured were different between auto-reactive and allo-reactive repertoires (Figure 6) and combined, showed 95% variance by two-way ANOVA ( $p \leq 0.0001$ ). In the IMN cohorts, auto-reactive CDR3 were more hydrophobic ( $p \leq 0.04$ ) and had higher mean grand average of hydropathicity index (GRAVY) scores showing these receptors had a low dielectric constant to water. Auto-reactive IMN CDR3 also had higher average molecular weights (Mw) ( $p = 0.008$ ) and average monoisotopic Mw ( $p \leq 0.008$ ), showing that these CDR3 were longer and/or have bulkier side chains. Auto-reactive IMN TCR also had higher mean theoretical pI suggesting that these receptors were more basic in charge. Interestingly, this pattern of hydrophobic, GRAVY, Mw and monoisotopic Mw metrics was inverted in the FLC system, showing that the physical-chemical properties of CDR3 loops are largely determined by antigen. Of note, was an increase in mean in theoretical pI in auto-reactive FLC-specific receptors which matched the IMN system.



**Figure 6.** Physical-chemical properties analysis of CDR3 loops between auto-reactive and allo-reactive TCRs. (A) IMN CDR3 (green) and (B) FLC CDR3 loops (orange) were compared by hydrophobicity, GRAVY, Mw average, Mw monoisotopic and theoretical pI scores. TCR were pooled and measured by appearance. \*  $p \leq 0.05$ , \*\*  $p \leq 0.01$ .

### 2.7. T-Cells Exhibit Divergent CDR3 Diversity Metrics between Auto-Reactive and Allo-Reactive Repertoires

We next quantitated and compared diversity metrics of IMN- and FLC-specific CDR3 loops from individual repertoires from auto-reactive and allo-reactive donors (Table 1). While individual diversity metrics did not reach significance, likely due to the low numbers of donors in each cohort ( $n = 3$ ), the combined metrics explained 45% ( $p \leq 0.0008$ ) and 50% ( $p \leq 0.0001$ ) of the variation between the IMN and FLC cohorts, respectively. Notably, the same  $\text{Log}_2$  fold-change ( $\text{Log}_2\text{FC}$ ) patterns were observed in both IMN and FLC systems suggesting these measures may be valid on an expanded donor analysis. The most heavily weighted diversity measurements are shown, as are the  $\text{Log}_2\text{FC}$  changes with increases shown in blue decreases shown in red.

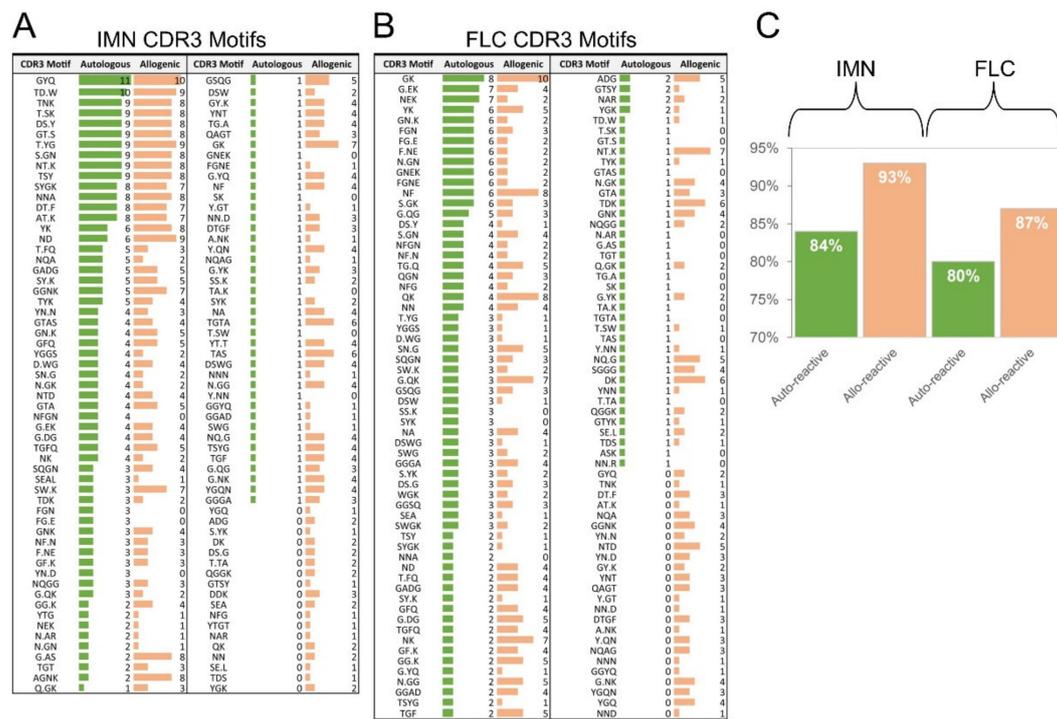
Using 12 formulas of diversity, we compared each system. Both IMN and FLC repertoires showed that allo-reactive TCR had increased numbers of unique clonotypes per individual (Menhinick index and Margalef Richness index), showed fewer multiple copy clonotypes but when they were present, were more prevalent in the repertoire (Dominance index and Gini coefficient). Allo-reactive TCR showed an overall increase in diversity which factors both total unique clonotypes and unique clonotypes with multiple copy numbers (Simpson's index, Reciprocal Simpson index, Berger-Parker Dominance index, and Inverted Berger-Parker Dominance index) and this diversity was more homogenous between individuals (Buzas and Gibson's index, Shannon index and Equitability index).

**Table 1.** (A), IMN T-cell repertoire; (B) FLC T-cell repertoire.

<b>A: IMN T-cell repertoire</b>		
Diversity Metric	<i>p</i> Value	Log <sub>2</sub> FC
Biodiversity	0.38	−0.15
Simpson's index	0.41	−5.45
Dominance index	0.41	0.51
Reciprocal Simpson index	0.75	−0.54
Shannon index	0.39	0.56
Menhinick index	0.37	0.57
Buzas and Gibson's index	0.35	0.34
Equitability index	0.40	0.43
Berger-Parker Dominance index	0.35	−2.41
Inverted Berger-Parker Dominance index	0.17	1.63
Margalef Richness index	0.38	0.59
Gini coefficient	0.34	−2.46
<b>B: FLC T-cell repertoire</b>		
Diversity Metric	<i>p</i> Value	Log <sub>2</sub> FC
Biodiversity	0.99	0.00
Simpson's index	0.73	−0.71
Dominance index	0.73	0.10
Reciprocal Simpson index	0.65	−0.65
Shannon index	0.90	0.05
Menhinick index	0.81	0.10
Buzas and Gibson's index	0.90	0.06
Equitability index	0.77	0.08
Berger-Parker Dominance index	0.84	−0.24
Inverted Berger-Parker Dominance index	0.56	0.81
Margalef Richness index	0.99	0.01
Gini coefficient	0.92	−0.14

### 2.8. T-Cells Exhibit Divergent CDR3 Motifs between Auto-Reactive and Allo-Reactive Repertoires

Finally, we examined CDR3 residue composition using software that detects common CDR3 motifs from a large database [43] (Figure 7). 118 CDR3 motifs were identified in the IMN repertoire and variation between the auto-reactive and allo-reactive repertoires was 84% by two-way ANOVA ( $p \leq 0.0001$ ). The allo-reactive repertoire contained more CDR3 motifs than the auto-reactive repertoire at 93% and 84%, respectively. 124 CDR3 motifs were identified in the FLC repertoire and variation between the auto-reactive and allo-reactive repertoires was 63% by two-way ANOVA ( $p \leq 0.0013$ ). Akin to the IMN system. The allo-reactive repertoire contained more CDR3 motifs than the auto-reactive repertoire at 87% and 80%, respectively. Interestingly, the YGQ motif was identified only in IMN and FLC allo-reactive repertoires. These data show that the TCR repertoires of IMN- and FLC-specific T-cells show different structural architecture between auto-reactive and allo-reactive individuals.



**Figure 7.** CDR3 motif analysis between auto-reactive and allo-reactive TCRs. (A) IMN and (B) FLC CDR3 loops found by motif searching. (C) Total percentage of IMN and FLC CDR3 loops found by motif searching. Auto-reactive (green) and allo-reactive (orange) sequences are shown as are the number of sequences encoding the motif.

### 3. Discussion

The distinct properties of allo-specific T-cell repertoires are not well known. This study used dextramers, single-cell TCR analysis and multiple annotation analysis to determine if alloreactive TCR repertoires could be identified from auto-reactive TCR repertoires. The enumeration of T-cells from ex vivo blood showed variability between antigens (IMN, FLC, YLE and KVA), with low mean frequencies, varying from 0.004 to 0.06% of CD8<sup>+</sup> T-cells even among auto-reactive donors. This was expected, as self-specific T-cells are typically low-frequency, apart from the abnormal MART-1 antigen which can have frequencies up to 0.35% of CD8<sup>+</sup> T-cells ex vivo [41], due to a unique germline focused docking mode [44,45]. Importantly, dextramer enumeration across all four antigens showed a consistent difference in mean frequency when comparing auto-reactive T-cells to allo-reactive T-cells.

Surprisingly, single-cell TCR analysis revealed multiple copies of unique clonotypes in some allo-reactive populations, suggesting possible cross-reactivity with a pre-existing memory T-cell response. However, it is unknown if these cells are genuine memory cells or high-frequency naïve cell clones which bear overproduced TCR from bias in recombination/selection [15,46–48] or bias in homeostatic proliferation [49]. Also of interest was the alignment of multiple copy number clonotypes per donor with mean dextramer frequency hierarchy. For instance, the IMN response showed more donors with multiple copies of unique clonotypes in the auto-reactive cohort, and the FLC response showed more donors with multiple copies of unique clonotypes in the allo-reactive cohort. This aligned with dominance hierarchies from dextramer scanning. However, this correlation is speculative and requires additional investigation across more donors, antigens, and alleles. None-the-less, divergent mean T-cell frequencies were consistently observed when comparing auto-reactive versus allo-reactive individuals.

Gene usage analysis of the IMN and FLC repertoires showed significant differences in TRBV and TRBJ profiles between cohorts with a consistent preference for TRBJ1-1, TRBV2-1 TRBV2-7 in both allo-reactive systems. Allo-reactive repertoires also showed increased TRAJ usage in both systems. Additionally, auto-reactive T-cell repertoires exhibited a

Gaussian distribution in CDR3 length whereas allo-reactive T-cell repertoires exhibited distorted CDR3 length distribution, with a tendency to encode shorter CDR3 loops. When examining TRBV/*JTRB*J and TRAV/TRAJ pairing between auto-reactive and allo-reactive repertoires, allo-reactive repertoires showed a much more complex pairing network across both systems.

Physical-chemical property analysis showed that auto-reactive CDR3 in the IMN system were more hydrophobic, which was supported by a corresponding trend in increased GRAVY scores. This is of interest given hydrophobic CDR3 loops promote the development of self-reactive T-cells [50] and are enriched in both mouse [50] and human [51] T regulatory ( $T_{reg}$ ) cells. This suggests that auto-reactive IMN T-cells may be of the  $T_{reg}$  lineage. However, phenotypic analysis is required to validate this observation. Allo-reactive IMN CDR3 loops showed significant decreases in average Mw and monoisotopic Mw. This made sense given the IMN repertoire was strongly skewed to short CDR3 sequences previously seen in gene analysis. The FLC system did not follow this divergence in hydrophobicity and Mw; however, multivariate analysis of physical-chemical scores could clearly distinguish the auto-reactive repertoire from the allo-reactive repertoire. Of interest, both IMN and FLC allo-reactive CDR3 loops showed lower mean theoretical pI scores, meaning allo-reactive receptors tended to be more acidic in charge.

Diversity metrics of the IMN and FLC systems could clearly distinguish auto-reactive repertoires from the allo-reactive repertoires in multivariate analysis but trends in univariate analysis. Of note, the same  $\text{Log}_2\text{FC}$  patterns were seen in both systems suggesting a possible general model for allo-reactive repertoires. Here, allo-reactive repertoires showed a general increase in diversity and consistency between individuals, meaning an increase in unique clonotypes, fewer unique multiple copy clonotypes and an increase in repertoire homogeneity between donors.

Motif analysis of CDR3 loops in the IMN and FLC systems could also clearly distinguish auto-reactive repertoires from the allo-reactive repertoires on multivariate analysis. Allo-reactive receptors in both systems showed increased motif pairing, showing that allo-reactive CDR3 loops exhibit more structural “options” when engaging an allo-reactive ligand. This finding aligns with increased TCR gene pairing observed in gene analysis and increased species observation in diversity analysis. While important, it must be said that these observations are preliminary, correlative and do not infer the cause of graft rejection. Future studies should be aimed at analysing additional sequences, additional antigens, additional donors for theory validation and patients for clinical validation.

Using single-cell sequencing, we have defined new auto-reactive and allo-reactive repertoires for the cancer antigens IMN and FLC, comprising 356 TCR $\alpha\beta$  pairs in the dataset. These receptors can be used for further exploratory research and have translatable potential. Indeed, foreign-antigen-specific TCR show up to a 10,000-fold increase in affinity compared with self-antigen-specific TCR [52] and are in the range of pathogen-specific TCRs ( $KD = 0.1\text{--}10 \mu\text{M}$ ) [53] which can bypass coreceptor help [54] and initiate cell death within five minutes [55]. Thus, the allo-reactive TCR identified here could apply to TCR gene transfer and adoptive cell therapy. Other potential therapies include “swapping” the allo-reactive peptide with an antagonist to prevent alloreactivity [54] or enhancing the affinity of the auto-reactive or allo-reactive TCR further using phage display [55]. These affinity enhanced receptors can be used as soluble drugs and are currently in clinical trials [56].

However, these receptors require functional validation before proceeding. Nevertheless, we are optimistic as the method used here has been published by our team numerous times, including functional validation in cancer-, Epstein-Barr virus-, cytomegalovirus- [39] and influenza-specific systems [57–59].

There are other variables to consider before TCR therapy, including “personal genetics” and “infectious history” which likely influence our personal T-cell repertoire makeup. Any two individuals can differ by up to 800 million single nucleotide polymorphisms (SNPs) between genomes [60]. This is termed personal genetics, and we currently do not

understand how our personal genomes fully impact immune function. Another variable is our infectivity record which is defined by which pathogens we acquire over the course of life and at what age, termed our infectious history [61], which can impact on future immune challenges.

In summary, we show allo-reactive T-cell repertoires and be defined from auto-reactive repertoires through differences in enumeration and TCR architecture when examining gene usage, gene pairing, CDR3 length, CDR3 physical-chemical properties, diversity measures and CDR3 motif appearance and usage. These data increase the knowledge base of the mechanisms underlying T-cell alloreactivity and may lead to new adoptive therapies, TCR therapies and the rational design of TCR for therapeutic use in autoimmunity, infectious disease, and cancer.

## 4. Materials and Methods

### 4.1. Sample Preparation and Processing

Healthy donor blood was studied using protocols carried out in accordance with guidelines and regulations under QIMR Berghofer Medical Research Institute (QIMRB) (HREC P2058, approved 4/9/2014). Informed consent was obtained from all participants in the study. The study was performed according to the rules of the Declaration of Helsinki of 1975. Peripheral blood mononuclear cells were separated, as previously described [62]. Briefly, peripheral blood mononuclear cells (PBMCs) were isolated by Ficoll-Paque PLUS (GE Healthcare, Chicago, IL, USA) density gradient centrifugation and cryopreserved in R10 medium (RPMI-1640 containing 10% FCS) supplemented with 10% DMSO (Sigma-Aldrich, St. Louis, MO, USA). PBMC were stored at  $10^6$  cells in a 1 mL vial and cryopreserved for batch analysis, which we have found increases the sensitivity of gene and protein detection and decreases experimental artefact. Eight-digit HLA typing was performed using NGS by BGI Genomics (BGI, Shanghai, China).

### 4.2. Flow Cytometry and Single-Cell Sorting

Cryopreserved PBMC were thawed in 37 °C R10 media with 1 uL/mL of DNase 1 (Thermo Fisher Scientific, Waltham, MA, USA) and 0.3 nM of dasatinib (Axon Medchem, Reston, VA, USA). APC-conjugated dextramer was added for 30 min on ice using either HLA-0201-IMNDMPHYM-APC (IMN), HLA-0201-FLCMKALLL-APC (FLC), HLA-A0201-KVAELVHFL-APC (KVA) or HLA-A0201-YLEPGPVTV-APC (YLE) dextramer (Immunodex, Toronto, ON, USA). We next surface stained the cells with CD14-Pacific Blue, CD16-Pacific Blue, CD19-Pacific blue (all BioLegend, San Diego, CA, USA), LIVE/DEAD<sup>®</sup> Fixable Aqua (Life Technologies, Carlsbad, CA, USA), CD3-PE and CD8-FITC (all BioLegend, San Diego, CA, USA) and sorting was performed as previously described [10]. Single cells were gated using FSC-W and FSC-H. Non-viable cells were dumped using LIVE/DEAD and CD14, CD16 and CD19. Dextramer<sup>+</sup> cells were sorted twice at low flow rates on a FACSAria IIu sorter (BD, Franklin Lakes, NJ, USA) into chilled 96-well twin-tec PCR plates (Eppendorf, Hamburg, Germany). The cells were then centrifuged at  $500 \times g$  for 2 min and immediately stored at  $-80$  °C.

### 4.3. Single-Cell TCR $\alpha\beta$ Amplification and Annotation

As previously described [63], cDNA was generated using VILO (Thermo Fisher, Waltham, MA, USA) according to the manufactures' protocol. A PCR was then performed to amplify the TCR locus followed by a then a nested PCR. Gels were run to confirm the amplification of positive products and cleaned for Sanger sequencing using QIAGEN plates (QIAGEN, Hilden, Germany). Post sequencing, TCR $\alpha$  and TCR $\beta$  transcripts were annotated using Bioedit and V-QUEST [64]. Random sampling and percentage usage graphs were generated in Excel 365 (Microsoft, 2019). Chord diagrams were generated in R using the *circlize* package [65]. CDR3 motifs were determined using GLIPH 2.0 [43]. Paired clones are listed in Supplementary Tables S1–S4.

#### 4.4. Diversity Analysis

T-cell repertoire diversity metrics were determined from total counts from individual donors. The calculations comprise:

$$\begin{array}{lll}
 \begin{array}{l} \text{Simpson's Index} \\ D' = \frac{\sum_i n_i (n_i - 1)}{N(N-1)} \end{array} & \begin{array}{l} \text{Dominance Index} \\ 1 - \frac{N(N-1)}{\sum_i n_i (n_i - 1)} \end{array} & \begin{array}{l} \text{Reciprocal} \\ \text{Simpson's Index} \\ 1 - D \end{array} \\
 \\
 \begin{array}{l} \text{Shannon's Index} \\ H' = \sum_i p_i \ln(p_i) \end{array} & \begin{array}{l} \text{Shannon-Weiner di-} \\ \text{versity index} \\ H' = -\sum_i \log_2 P_i \end{array} & \begin{array}{l} \text{Margalef Rich-} \\ \text{ness Index} \\ \frac{n_{max}}{N} \end{array} \\
 \\
 \begin{array}{l} \text{Menhenik Index} \\ \frac{S}{\sqrt{\sum_i n_i}} \end{array} & \begin{array}{l} \text{Buzas and Gibson's} \\ \text{Index} \\ e^{-\frac{\sum_i (n_i) \ln(n_i)}{S}} \end{array} & \begin{array}{l} \text{Inverted Berger-} \\ \text{Parker Dominance} \\ \text{Index} \\ \frac{N}{n_{max}} \end{array} \\
 \\
 \begin{array}{l} \text{Equitability Index} \\ \frac{\sum_i \frac{n_i}{N} - \ln \left( \frac{n_i}{N} \right)}{\ln N} \end{array} & \begin{array}{l} \text{Berger-Parker} \\ \text{Dominance Index} \\ \frac{n_i}{N} \end{array} & \begin{array}{l} \text{Gini Coefficient} \\ \frac{2\sum_i i n_i - N + 1}{n \sum_i n_i} \frac{N + 1}{N} \end{array}
 \end{array}$$

#### 4.5. Physical-Chemical Properties

CDR3 loop hydrophobicity, GRAVY, average Mw, monoisotopic Mw and theoretical pI were determined using the Proteotypic Peptide Analyzing Tool (Thermo Fisher Scientific, Waltham, MA, USA).

#### 4.6. Statistical Analysis

TCR repertoire annotation was performed using IMGT/V-QUEST [64] and Excel 365 (Microsoft, USA). Post TCR gene alignment and CDR3 annotation, repertoires were then binned into individual dextramer<sup>+</sup> repertoires or total pooled repertoires. Depending on the analysis, clonotypes were assigned to total repertoires (auto-reactive versus allo-reactive), abundance repertoires (where multiple clonal sequences were included) and appearance repertoires (where clonal sequences were decreased to one count and clonotypes randomly selected). Data was first checked for normality and lognormality using the D'Agostino-Pearson omnibus normality test and or the Shapiro-Wilk normality test. If the data failed these tests, a one sample t-test was performed with a Wilcoxon signed-rank test with mean calculated. Post quality control, a Mann-Whitney rank test, two-way ANOVA with Tukey's multiple comparison test was performed using Prism V9.01 (Graphpad, USA). The *p*-value cut-offs are defined by \* *p* < 0.05, \*\* *p* < 0.01, \*\*\* *p* < 0.001, \*\*\*\* *p* < 0.0001.

**Supplementary Materials:** The following are available online at <https://www.mdpi.com/1422-0067/22/4/1625/s1>.

**Author Contributions:** Conceptualisation, J.J.M., methodology and analysis, O.L.H., E.J.G., K.K., M.A.F. and J.J.M., manuscript preparation O.L.H., E.J.G., T.H.O.N., K.K., M.A.F. and J.J.M., supervision J.J.M. funding acquisition; J.J.M. Funding for this project was supported by the Prostate Cancer Foundation Australia (PCFA) (NCG 3913). All authors have read and agreed to the published version of the manuscript.

**Funding:** This project was supported by the Prostate Cancer Foundation Australia (PCFA) Research Grant (NCG 3913). E.J.G. was supported by a National and Medical Research Council of Australia (NHMRC) C.J Martin Fellowship (1110429) and an Australian Research Council DECRA Fellowship (DE210101479), K.K. is supported by an NHMRC Leadership Investigator Grant (1173871), M.A.F. is supported by National and Medical Research Council of Australia (NHMRC) fellowship (5121190) and J.J.M. was supported by CDF Level 1 & 2 NHMRC Fellowships (1031652 & 1131732).

**Institutional Review Board Statement:** This study used protocols carried out in accordance with guidelines and regulations under QIMR Berghofer Medical Research Institute (QIMRB) (HREC P2058, approved 4/9/2014).

**Informed Consent Statement:** Informed consent was obtained from all participants in the study. The study was performed according to the rules of the Declaration of Helsinki of 1975.

**Data Availability Statement:** The TCR data is available at <https://vdjdb.cdr3.net>.

**Conflicts of Interest:** J.J.M. reports receiving a commercial research grant from Brandan Capitol, OneVentures and AbbVie. There is no scientific overlap.

## Abbreviations

TCR	T-cell receptor
MHC	Major histocompatibility complex
HLA	Human leukocyte antigen complex
p	Peptide within the MHC/HLA binding groove
TRBV	T-cell receptor $\beta$ gene
TRBJ	T-cell receptor $\beta$ joining gene
TRBD	T-cell receptor $\beta$ diversity gene
TRAV	T-cell receptor $\alpha$ region gene
TRAJ	T-cell receptor $\alpha$ joining gene
CDR	Complementarity determining region

## References

- Felix, N.J.; Allen, P.M. Specificity of T-cell alloreactivity. *Nat. Rev. Immunol.* **2007**, *7*, 942–953. [[CrossRef](#)]
- Gkazi, A.S.; Margetts, B.K.; Attenborough, T.; Mhaldien, L.; Standing, J.F.; Oakes, T.; Heather, J.M.; Booth, J.; Pasquet, M.; Chiesa, R.; et al. Clinical T Cell Receptor Repertoire Deep Sequencing and Analysis: An Application to Monitor Immune Reconstitution Following Cord Blood Transplantation. *Front. Immunol.* **2018**, *9*, 2547. [[CrossRef](#)]
- Klarenbeek, P.L.; Remmerswaal, E.B.; ten Berge, I.J.; Doorenspleet, M.E.; van Schaik, B.D.; Esveldt, R.E.; Koch, S.D.; ten Brinke, A.; van Kampen, A.H.; Bemelman, F.J.; et al. Deep sequencing of antiviral T-cell responses to HCMV and EBV in humans reveals a stable repertoire that is maintained for many years. *PLoS Pathog.* **2012**, *8*, e1002889. [[CrossRef](#)] [[PubMed](#)]
- van Heijst, J.W.; Ceberio, I.; Lipuma, L.B.; Samilo, D.W.; Wasilewski, G.D.; Gonzales, A.M.; Nieves, J.L.; van den Brink, M.R.; Perales, M.A.; Pamer, E.G. Quantitative assessment of T cell repertoire recovery after hematopoietic stem cell transplantation. *Nat. Med.* **2013**, *19*, 372–377. [[CrossRef](#)] [[PubMed](#)]
- Miller, M.L.; McIntosh, C.M.; Williams, J.B.; Wang, Y.; Hollinger, M.K.; Isaad, N.J.; Moon, J.J.; Gajewski, T.F.; Chong, A.S.; Alegre, M.L. Distinct Graft-Specific TCR Avidity Profiles during Acute Rejection and Tolerance. *Cell Rep.* **2018**, *24*, 2112–2126. [[CrossRef](#)] [[PubMed](#)]
- Nikolich-Zugich, J. High specificity, not degeneracy, allows T cell alloresponses. *Nat. Immunol.* **2007**, *8*, 335–337. [[CrossRef](#)]
- Archbold, J.K.; Macdonald, W.A.; Miles, J.J.; Brennan, R.M.; Kjer-Nielsen, L.; McCluskey, J.; Burrows, S.R.; Rossjohn, J. Alloreactivity between disparate cognate and allogeneic pMHC-I complexes is the result of highly focused, peptide-dependent structural mimicry. *J. Biol. Chem.* **2006**, *281*, 34324–34332. [[CrossRef](#)]
- Liu, Y.C.; Miles, J.J.; Neller, M.A.; Gostick, E.; Price, D.A.; Purcell, A.W.; McCluskey, J.; Burrows, S.R.; Rossjohn, J.; Gras, S. Highly divergent T-cell receptor binding modes underlie specific recognition of a bulged viral peptide bound to a human leukocyte antigen class I molecule. *J. Biol. Chem.* **2013**, *288*, 15442–15454. [[CrossRef](#)]
- Rossjohn, J.; Gras, S.; Miles, J.J.; Turner, S.J.; Godfrey, D.I.; McCluskey, J. T cell antigen receptor recognition of antigen-presenting molecules. *Annu. Rev. Immunol.* **2015**, *33*, 169–200. [[CrossRef](#)]
- Miles, J.J.; McCluskey, J.; Rossjohn, J.; Gras, S. Understanding the complexity and malleability of T-cell recognition. *Immunol. Cell Biol.* **2015**, *93*, 433–441. [[CrossRef](#)]
- Chapman, J.R.; Webster, A.C.; Wong, G. Cancer in the transplant recipient. *Cold Spring Harb. Perspect. Med.* **2013**, *3*. [[CrossRef](#)]
- Benichou, G.; Gonzalez, B.; Marino, J.; Ayasoufi, K.; Valujskikh, A. Role of Memory T Cells in Allograft Rejection and Tolerance. *Front. Immunol.* **2017**, *8*, 170. [[CrossRef](#)]
- Ingulli, E. Mechanism of cellular rejection in transplantation. *Pediatr. Nephrol.* **2010**, *25*, 61–74. [[CrossRef](#)]
- DeWolf, S.; Sykes, M. Alloimmune T cells in transplantation. *J. Clin. Investig.* **2017**, *127*, 2473–2481. [[CrossRef](#)]
- Watkins, T.S.; Miles, J.J. The human T cell receptor repertoire in health and disease and potential for omics integration. *Immunol. Cell Biol.* **2020**. [[CrossRef](#)]
- Kjer-Nielsen, L.; Clements, C.S.; Purcell, A.W.; Brooks, A.G.; Whisstock, J.C.; Burrows, S.R.; McCluskey, J.; Rossjohn, J. A structural basis for the selection of dominant alphabeta T cell receptors in antiviral immunity. *Immunity* **2003**, *18*, 53–64. [[CrossRef](#)]

17. Macdonald, W.A.; Chen, Z.; Gras, S.; Archbold, J.K.; Tynan, F.E.; Clements, C.S.; Bharadwaj, M.; Kjer-Nielsen, L.; Saunders, P.M.; Wilce, M.C.; et al. T cell allorecognition via molecular mimicry. *Immunity* **2009**, *31*, 897–908. [[CrossRef](#)] [[PubMed](#)]
18. Gras, S.; Burrows, S.R.; Kjer-Nielsen, L.; Clements, C.S.; Liu, Y.C.; Sullivan, L.C.; Bell, M.J.; Brooks, A.G.; Purcell, A.W.; McCluskey, J.; et al. The shaping of T cell receptor recognition by self-tolerance. *Immunity* **2009**, *30*, 193–203. [[CrossRef](#)] [[PubMed](#)]
19. Speir, J.A.; Garcia, K.C.; Brunmark, A.; Degano, M.; Peterson, P.A.; Teyton, L.; Wilson, I.A. Structural basis of 2C TCR allorecognition of H-2Ld peptide complexes. *Immunity* **1998**, *8*, 553–562. [[CrossRef](#)]
20. Reiser, J.B.; Darnault, C.; Guimezanes, A.; Gregoire, C.; Mosser, T.; Schmitt-Verhulst, A.M.; Fontecilla-Camps, J.C.; Malissen, B.; Housset, D.; Mazza, G. Crystal structure of a T cell receptor bound to an allogeneic MHC molecule. *Nat. Immunol.* **2000**, *1*, 291–297. [[CrossRef](#)]
21. Colf, L.A.; Bankovich, A.J.; Hanick, N.A.; Bowerman, N.A.; Jones, L.L.; Kranz, D.M.; Garcia, K.C. How a single T cell receptor recognizes both self and foreign MHC. *Cell* **2007**, *129*, 135–146. [[CrossRef](#)]
22. Jones, L.L.; Colf, L.A.; Stone, J.D.; Garcia, K.C.; Kranz, D.M. Distinct CDR3 conformations in TCRs determine the level of cross-reactivity for diverse antigens, but not the docking orientation. *J. Immunol.* **2008**, *181*, 6255–6264. [[CrossRef](#)]
23. Wang, Y.; Singh, N.K.; Spear, T.T.; Hellman, L.M.; Piepenbrink, K.H.; McMahan, R.H.; Rosen, H.R.; Vander Kooi, C.W.; Nishimura, M.I.; Baker, B.M. How an allreactive T-cell receptor achieves peptide and MHC specificity. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, E4792–E4801. [[CrossRef](#)]
24. Yin, L.; Huseby, E.; Scott-Browne, J.; Rubtsova, K.; Pinilla, C.; Crawford, F.; Marrack, P.; Dai, S.; Kappler, J.W. A Single T Cell Receptor Bound to Major Histocompatibility Complex Class I and Class II Glycoproteins Reveals Switchable TCR Conformers. *Immunity* **2011**, *35*, 23–33. [[CrossRef](#)] [[PubMed](#)]
25. Rist, M.; Smith, C.; Bell, M.J.; Burrows, S.R.; Khanna, R. Cross-recognition of HLA DR4 alloantigen by virus-specific CD8+ T cells: A new paradigm for self-/nonself-recognition. *Blood* **2009**, *114*, 2244–2253. [[CrossRef](#)] [[PubMed](#)]
26. Rist, M.J.; Theodosis, A.; Croft, N.P.; Neller, M.A.; Welland, A.; Chen, Z.; Sullivan, L.C.; Burrows, J.M.; Miles, J.J.; Brennan, R.M.; et al. HLA peptide length preferences control CD8+ T cell responses. *J. Immunol.* **2013**, *191*, 561–571. [[CrossRef](#)]
27. Rist, M.J.; Hibbert, K.M.; Croft, N.P.; Smith, C.; Neller, M.A.; Burrows, J.M.; Miles, J.J.; Purcell, A.W.; Rossjohn, J.; Gras, S.; et al. T Cell Cross-Reactivity between a Highly Immunogenic EBV Epitope and a Self-Peptide Naturally Presented by HLA-B\*18:01+ Cells. *J. Immunol.* **2015**, *194*, 4668–4675. [[CrossRef](#)]
28. Cole, D.K.; Miles, K.M.; Madura, F.; Holland, C.J.; Schauenburg, A.J.; Godkin, A.J.; Bulek, A.M.; Fuller, A.; Akpovwa, H.J.; Pymm, P.G.; et al. T-cell receptor (TCR)-peptide specificity overrides affinity-enhancing TCR-major histocompatibility complex interactions. *J. Biol. Chem.* **2014**, *289*, 628–638. [[CrossRef](#)] [[PubMed](#)]
29. Alves, P.M.; Faure, O.; Graff-Dubois, S.; Gross, D.A.; Cornet, S.; Chouaib, S.; Miconnet, I.; Lemonnier, F.A.; Kosmatopoulos, K. EphA2 as target of anticancer immunotherapy: Identification of HLA-A\*0201-restricted epitopes. *Cancer Res.* **2003**, *63*, 8476–8480.
30. Marrari, A.; Iero, M.; Pilla, L.; Villa, S.; Salvioni, R.; Valdagni, R.; Parmiani, G.; Rivoltini, L. Vaccination therapy in prostate cancer. *Cancer Immunol. Immunother.* **2007**, *56*, 429–445. [[CrossRef](#)]
31. Novellino, L.; Castelli, C.; Parmiani, G. A listing of human tumor antigens recognized by T cells: March 2004 update. *Cancer Immunol. Immunother.* **2005**, *54*, 187–207. [[CrossRef](#)]
32. McNeel, D.G.; Olsen, B.M. Prostate Cancer Vaccine. U.S. Patent 9,433,668, B2, 2016.
33. Wang, L.G.; Johnson, E.M.; Kinoshita, Y.; Babb, J.S.; Buckley, M.T.; Liebes, L.F.; Melamed, J.; Liu, X.M.; Kurek, R.; Ossowski, L.; et al. Androgen receptor overexpression in prostate cancer linked to Pur alpha loss from a novel repressor complex. *Cancer Res.* **2008**, *68*, 2678–2688. [[CrossRef](#)]
34. Giovannelli, P.; Di Donato, M.; Galasso, G.; Di Zazzo, E.; Bilancio, A.; Migliaccio, A. The Androgen Receptor in Breast Cancer. *Front. Endocrinol. (Lausanne)* **2018**, *9*, 492. [[CrossRef](#)]
35. Zhang, H.; Li, X.X.; Yang, Y.; Zhang, Y.; Wang, H.Y.; Zheng, X.F.S. Significance and mechanism of androgen receptor overexpression and androgen receptor/mechanistic target of rapamycin cross-talk in hepatocellular carcinoma. *Hepatology* **2018**, *67*, 2271–2286. [[CrossRef](#)]
36. Weon, J.L.; Potts, P.R. The MAGE protein family and cancer. *Curr. Opin. Cell Biol.* **2015**, *37*, 1–8. [[CrossRef](#)]
37. Zhang, Y.; Zhang, Y.; Zhang, L. Expression of cancer-testis antigens in esophageal cancer and their progress in immunotherapy. *J. Cancer Res. Clin. Oncol.* **2019**, *145*, 281–291. [[CrossRef](#)] [[PubMed](#)]
38. Fossa, A.; Alsoe, L.; Cramer, R.; Funderud, S.; Gaudernack, G.; Smeland, E.B. Serological cloning of cancer/testis antigens expressed in prostate cancer using cDNA phage surface display. *Cancer Immunol. Immunother.* **2004**, *53*, 431–438. [[CrossRef](#)] [[PubMed](#)]
39. Neller, M.A.; Ladell, K.; McLaren, J.E.; Matthews, K.K.; Gostick, E.; Pentier, J.M.; Dolton, G.; Schauenburg, A.J.; Koning, D.; Fontaine Costa, A.I.; et al. Naive CD8(+) T-cell precursors display structured TCR repertoires and composite antigen-driven selection dynamics. *Immunol. Cell Biol.* **2015**, *93*, 625–633. [[CrossRef](#)] [[PubMed](#)]
40. Dolton, G.; Tungatt, K.; Lloyd, A.; Bianchi, V.; Theaker, S.M.; Trimby, A.; Holland, C.J.; Donia, M.; Godkin, A.J.; Cole, D.K.; et al. More tricks with tetramers: A practical guide to staining T cells with peptide-MHC multimers. *Immunology* **2015**, *146*, 11–22. [[CrossRef](#)]
41. Dolton, G.; Lissina, A.; Skowera, A.; Ladell, K.; Tungatt, K.; Jones, E.; Kronenberg-Versteeg, D.; Akpovwa, H.; Pentier, J.M.; Holland, C.J.; et al. Comparison of peptide-major histocompatibility complex tetramers and dextramers for the identification of antigen-specific T cells. *Clin. Exp. Immunol.* **2014**, *177*, 47–63. [[CrossRef](#)]

42. Lissina, A.; Ladell, K.; Skowera, A.; Clement, M.; Edwards, E.; Seggewiss, R.; van den Berg, H.A.; Gostick, E.; Gallagher, K.; Jones, E.; et al. Protein kinase inhibitors substantially improve the physical detection of T-cells with peptide-MHC tetramers. *J. Immunol. Methods* **2009**, *340*, 11–24. [[CrossRef](#)] [[PubMed](#)]
43. Glanville, J.; Huang, H.; Nau, A.; Hatton, O.; Wagar, L.E.; Rubelt, F.; Ji, X.; Han, A.; Krams, S.M.; Pettus, C.; et al. Identifying specificity groups in the T cell receptor repertoire. *Nature* **2017**, *547*, 94–98. [[CrossRef](#)] [[PubMed](#)]
44. Cole, D.K.; Yuan, F.; Rizkallah, P.J.; Miles, J.J.; Gostick, E.; Price, D.A.; Gao, G.F.; Jakobsen, B.K.; Sewell, A.K. Germ line-governed recognition of a cancer epitope by an immunodominant human T-cell receptor. *J. Biol. Chem.* **2009**, *284*, 27281–27289. [[CrossRef](#)]
45. Madura, F.; Rizkallah, P.J.; Legut, M.; Holland, C.J.; Fuller, A.; Bulek, A.; Schauenburg, A.J.; Trimby, A.; Hopkins, J.R.; Wells, S.A.; et al. TCR-induced alteration of primary MHC peptide anchor residue. *Eur. J. Immunol.* **2019**, *49*, 1052–1066. [[CrossRef](#)] [[PubMed](#)]
46. Miles, J.J.; Douek, D.C.; Price, D.A. Bias in the alphabeta T-cell repertoire: Implications for disease pathogenesis and vaccination. *Immunol. Cell Biol.* **2011**, *89*, 375–387. [[CrossRef](#)]
47. Venturi, V.; Price, D.A.; Douek, D.C.; Davenport, M.P. The molecular basis for public T-cell responses? *Nat. Rev. Immunol.* **2008**, *8*, 231–238. [[CrossRef](#)]
48. Melenhorst, J.J.; Lay, M.D.; Price, D.A.; Adams, S.D.; Zeilah, J.; Sosa, E.; Hensel, N.F.; Follmann, D.; Douek, D.C.; Davenport, M.P.; et al. Contribution of TCR-beta locus and HLA to the shape of the mature human Vbeta repertoire. *J. Immunol.* **2008**, *180*, 6484–6489. [[CrossRef](#)]
49. Min, B. Spontaneous T Cell Proliferation: A Physiologic Process to Create and Maintain Homeostatic Balance and Diversity of the Immune System. *Front. Immunol.* **2018**, *9*, 547. [[CrossRef](#)]
50. Stadinski, B.D.; Shekhar, K.; Gomez-Tourino, I.; Jung, J.; Sasaki, K.; Sewell, A.K.; Peakman, M.; Chakraborty, A.K.; Huseby, E.S. Hydrophobic CDR3 residues promote the development of self-reactive T cells. *Nat. Immunol.* **2016**, *17*, 946–955. [[CrossRef](#)] [[PubMed](#)]
51. Henderson, L.A.; Volpi, S.; Frugoni, F.; Janssen, E.; Kim, S.; Sundel, R.P.; Dedeoglu, F.; Lo, M.S.; Hazen, M.M.; Beth Son, M.; et al. Next-Generation Sequencing Reveals Restriction and Clonotypic Expansion of Treg Cells in Juvenile Idiopathic Arthritis. *Arthritis Rheumatol.* **2016**, *68*, 1758–1768. [[CrossRef](#)]
52. Rosenthal, K.M.; Edwards, L.J.; Sabatino, J.J., Jr.; Hood, J.D.; Wasserman, H.A.; Zhu, C.; Evavold, B.D. Low 2-dimensional CD4 T cell receptor affinity for myelin sets in motion delayed response kinetics. *PLoS ONE* **2012**, *7*, e32562. [[CrossRef](#)] [[PubMed](#)]
53. Miles, J.J.; Bulek, A.M.; Cole, D.K.; Gostick, E.; Schauenburg, A.J.; Dolton, G.; Venturi, V.; Davenport, M.P.; Tan, M.P.; Burrows, S.R.; et al. Genetic and structural basis for selection of a ubiquitous T cell receptor deployed in Epstein-Barr virus infection. *PLoS Pathog.* **2010**, *6*, e1001198. [[CrossRef](#)] [[PubMed](#)]
54. Ely, L.K.; Green, K.J.; Beddoe, T.; Clements, C.S.; Miles, J.J.; Bottomley, S.P.; Zernich, D.; Kjer-Nielsen, L.; Purcell, A.W.; McCluskey, J.; et al. Antagonism of antiviral and allogeneic activity of a human public CTL clonotype by a single altered peptide ligand: Implications for allograft rejection. *J. Immunol.* **2005**, *174*, 5593–5601. [[CrossRef](#)] [[PubMed](#)]
55. Miles, J.J.; Silins, S.L.; Burrows, S.R. Engineered T cell receptors and their potential in molecular medicine. *Curr. Med. Chem.* **2006**, *13*, 2725–2736. [[CrossRef](#)]
56. Boudousquie, C.; Bossi, G.; Hurst, J.M.; Rygiel, K.A.; Jakobsen, B.K.; Hassan, N.J. Polyfunctional response by ImmTAC (IMCgp100) redirected CD8(+) and CD4(+) T cells. *Immunology* **2017**, *152*, 425–438. [[CrossRef](#)]
57. Sant, S.; Grzelak, L.; Wang, Z.; Pizzolla, A.; Koutsakos, M.; Crowe, J.; Loudovaris, T.; Mannering, S.I.; Westall, G.P.; Wakim, L.M.; et al. Single-Cell Approach to Influenza-Specific CD8(+) T Cell Receptor Repertoires Across Different Age Groups, Tissues, and Following Influenza Virus Infection. *Front. Immunol.* **2018**, *9*, 1453. [[CrossRef](#)]
58. Van de Sandt, C.E.; Clemens, E.B.; Grant, E.J.; Rowntree, L.C.; Sant, S.; Halim, H.; Crowe, J.; Cheng, A.C.; Kotsimbos, T.C.; Richards, M.; et al. Challenging immunodominance of influenza-specific CD8(+) T cell responses restricted by the risk-associated HLA-A\*68:01 allomorph. *Nat. Commun.* **2019**, *10*, 5579. [[CrossRef](#)]
59. Kedzierska, K.; Venturi, V.; Field, K.; Davenport, M.P.; Turner, S.J.; Doherty, P.C. Early establishment of diverse T cell receptor profiles for influenza-specific CD8(+)CD62L(hi) memory T cells. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 9184–9189. [[CrossRef](#)]
60. Ma, Z.S.; Li, L.; Zhang, Y.P. Defining Individual-Level Genetic Diversity and Similarity Profiles. *Sci. Rep.* **2020**, *10*, 5805. [[CrossRef](#)]
61. Trowsdale, J. The MHC, disease and selection. *Immunol. Lett.* **2011**, *137*, 1–8. [[CrossRef](#)]
62. Lepletier, A.; Lutzky, V.P.; Mittal, D.; Stannard, K.; Watkins, T.S.; Ratnatunga, C.N.; Smith, C.; McGuire, H.M.; Kemp, R.A.; Mukhopadhyay, P.; et al. The immune checkpoint CD96 defines a distinct lymphocyte phenotype and is highly expressed on tumor-infiltrating T cells. *Immunol. Cell Biol.* **2019**, *97*, 152–164. [[CrossRef](#)] [[PubMed](#)]
63. Wang, G.C.; Dash, P.; McCullers, J.A.; Doherty, P.C.; Thomas, P.G. T cell receptor alphabeta diversity inversely correlates with pathogen-specific antibody levels in human cytomegalovirus infection. *Sci. Transl. Med.* **2012**, *4*, 128ra142. [[CrossRef](#)]
64. Li, S.; Lefranc, M.P.; Miles, J.J.; Alamyar, E.; Giudicelli, V.; Duroux, P.; Freeman, J.D.; Corbin, V.D.; Scheerlinck, J.P.; Frohman, M.A.; et al. IMGT/HighV QUEST paradigm for T cell receptor IMGT clonotype diversity and next generation repertoire immunoprofiling. *Nat. Commun.* **2013**, *4*, 2333. [[CrossRef](#)] [[PubMed](#)]
65. Gu, Z.; Gu, L.; Eils, R.; Schlesner, M.; Brors, B. Circlize Implements and enhances circular visualization in R. *Bioinformatics* **2014**, *30*, 2811–2812. [[CrossRef](#)] [[PubMed](#)]