T-Cell Receptor Repertoire Sequencing and Its Applications

Subjects: Immunology

Contributor: Lucia Mazzotti , Anna Gaimari , Sara Bravaccini , Roberta Maltoni , Claudio Cerchione , Manel Juan , Europa Azucena-Gonzalez Navarro , Anna Pasetto , Daniela Nascimento Silva , Valentina Ancarani , Vittorio Sambri , Luana Calabrò , Giovanni Martinelli , Massimiliano Mazza

The immune system is a dynamic feature of each individual and a footprint of our unique internal and external exposures. Indeed, the type and level of exposure to physical and biological agents shape the development and behavior of this complex and diffuse system. Many pathological conditions depend on how our immune system responds or does not respond to a pathogen or a disease or on how the regulation of immunity is altered by the disease itself. T-cells are important players in adaptive immunity and, together with B-cells, define specificity and monitor the internal and external signals that our organism perceives through its specific receptors, TCRs and BCRs, respectively. Today, high-throughput sequencing (HTS) applied to the TCR repertoire has opened a window of opportunity to disclose T-cell repertoire development and behavior down to the clonal level. Although TCR repertoire sequencing is easily accessible today, it is important to deeply understand the available technologies for choosing the best fit for the specific experimental needs and questions. Here, an updated overview of TCR repertoire sequencing strategies, providers and applications to infectious diseases and cancer to guide researchers' choice through the multitude of available options is provided. The possibility of extending the TCR repertoire to HLA characterization will be of pivotal importance in the near future to understand how specific HLA genes shape T-cell responses in different pathological contexts and will add a level of comprehension that was unthinkable just a few years ago.

TCR repertoire TCR sequencing infectious diseases cancer immunotherapy HLA TILs COVID-19 T-cell response

1. Introduction

The immune system guards multicellular living organisms from the damage of pathogens [1].

Adaptive immunity defense is elicited by large numbers of T-cell receptors (TCRs) and B-cell receptors (BCRs) [1], and perception and adaptation to external insults enhance pre-existing and generate de novo receptors that the immune system records and retains as immunological memory (e.g., with vaccination) [2].

T-cells, the main actors in cytotoxic immune responses and also contributing to the full activation of the humoral adaptive response, include mainly αβ T-lymphocytes expressing TCR alpha and beta chains and γδ T-cells, a

smaller population expressing TCR gamma and delta chains; in humans, these different types of T-cells are about 95% and 1–5%, respectively [1][3][4].

The αβ T-cells are essential in cellular immunity since they mediate recognition of antigen peptides in a major histocompatibility complex class I and II (MHC) restricted manner, defining antigen complexes and driving the antigen-specific adaptive immune response against non-self-perceived antigens, including pathogens and cancer neo-antigens, through the TCRs [5].

The TCR is a heterodimeric plasma membrane protein located on T-cells, and it is composed of two paired chains [6][7] whose loci are organized as gene segment families comprising a variable (V) gene, a diversity (D) gene, a joining (J) gene and a constant (C) gene.

During T-cell maturation in the thymus, the β and δ chains' VDJ gene fragments are rearranged as one allele of each gene segment randomly recombines with the others, composing a functional V segment [8], while, for α and γ chains only, VJ fragments recombine as there is no D gene [4]. Each coding segment is flanked by conserved DNA regions, the recombination signal sequences (RSSs), which serve as cleavage signals for enzymes encoded by recombination activating genes RAG1 and RAG2. After RSSs recognition, these enzymes introduce double strand breaks (DSBs), initiating the V(D)J recombination process [9][10].

The V region of the α and β chains has three hypervariable complementary-determining regions (CDRs): CDR1 and CDR2 come with the V gene and play a role in both the interaction and stabilization of the TCR-MHC complex [8], while the most variable CDR3 is encoded by VDJ or VJ segments [3] and determines most of the binding specificity of the TCRs to the antigen-MHC complex [8][11].

In addition to somatic rearrangements (recombination diversity), insertions or deletions may occur randomly at the junctions of both VD and DJ fragments (junctional diversity) for a functional TCR sequence [3][12]. As the recombination mechanism just described occurs separately for both chains, the combinatorial diversity increases even further the number of possible TCR variants [13] (**Table 1**).

Table 1. Mechanisms involved in the generation of TCR diversity.

Sources of α/β TCR Diversity	
Recombination of the T-cell α-genes on chromosome 14 and T-cell β-genes on chromosome 7 by RAG1/2 enzymes.	Total α-genes combination: 2392 Total β-genes combination: 1248 Total αβ-genes combination: 2,985,216 [6].
Theoretical diversity by pairing of different in-frame α- and β-chains plus junctional diversity by terminal deoxynucleotidyl transferase activity [6].	10^{15}–10^{61} [14]

Sources of α/β TCR Diversity		
Experimental diversity evaluation by deep sequencing. [4]	[15]	10^4–10^6, based on the amount of the sample.

ding by T-
ry, T-cells
onotypes)
that share the same TCR sequences conserved during mature T-lymphocytes mitosis [15].

TCRβ clonotypes in individuals are estimated to 10^6–10^8 [16][17] on a total estimated number of 10^{12} circulating lymphocytes [18]. Therefore, even if the actual diversity of the paired TCRα and TRBβ repertoire is still debated and estimated to range from a lower limit of 10^{15} up to 10^{61} [14], is it possible to assay an overall complexity of 10^4–10^6 T-cells only in an individual experiment [19]? At present, no technologies are available and sufficiently empowered to describe such an immense source of variability.

The entire amazingly broad and diverse assembly of TCR sequences, known as the TCR repertoire, is composed of naive T-cells, antigen-inexperienced cells shaped in the thymus, where new antigen specificities are introduced [20][21], and memory T-lymphocytes, antigen-experienced cells whose specificity depends on antigen exposure through people's lives and persisting long-term [22].

The TCR-MHC affinity can delineate the T-cells propensity, defining the level of equilibrium between effector and memory T-cells [23][24].

Naive and memory T-cells make up a unique TCR footprint that is different even in genetically identical twins [25] due to the intrinsically stochastic nature of TCR generation and immune system personal experience.

Moreover, the TCR repertoire footprint dynamically evolves according to the challenges with which the immune system is confronted [4], such as infections, aging, autoimmune diseases, cancer and many other stimuli. Importantly, the exposure to an antigen triggers a massive expansion of antigen-specific T-cells, altering the composition of the TCR repertoire in favor of specific clonotypes (TCR bias) [26][27] that can be correlated to a biological process and exploited as "molecular barcodes" both in health and diseases [15].

2. TCR Repertoire Analysis

It allows the definition of these types of "barcodes" in different contexts, revealing, for example, important information about a successful antitumor T-cell response, on how to improve efficacy and safety of immune checkpoint inhibitors (ICI), on the tumor microenvironment (TME) characterization, on the immune response during disease development and treatment, on minimal residual disease (MRD) assessment, transplantation, autoimmune disease and infectious disease characterization and therapy [1][4][8][13][28][29][30].

In this entry, researchers describe TCR repertoire sequencing strategies and applications (**Figure 1**) in individuals exposed to infectious agents, such as HIV, HBV, HCV and SARS-CoV-2, and to cancer, with an updated overview of the available technologies.

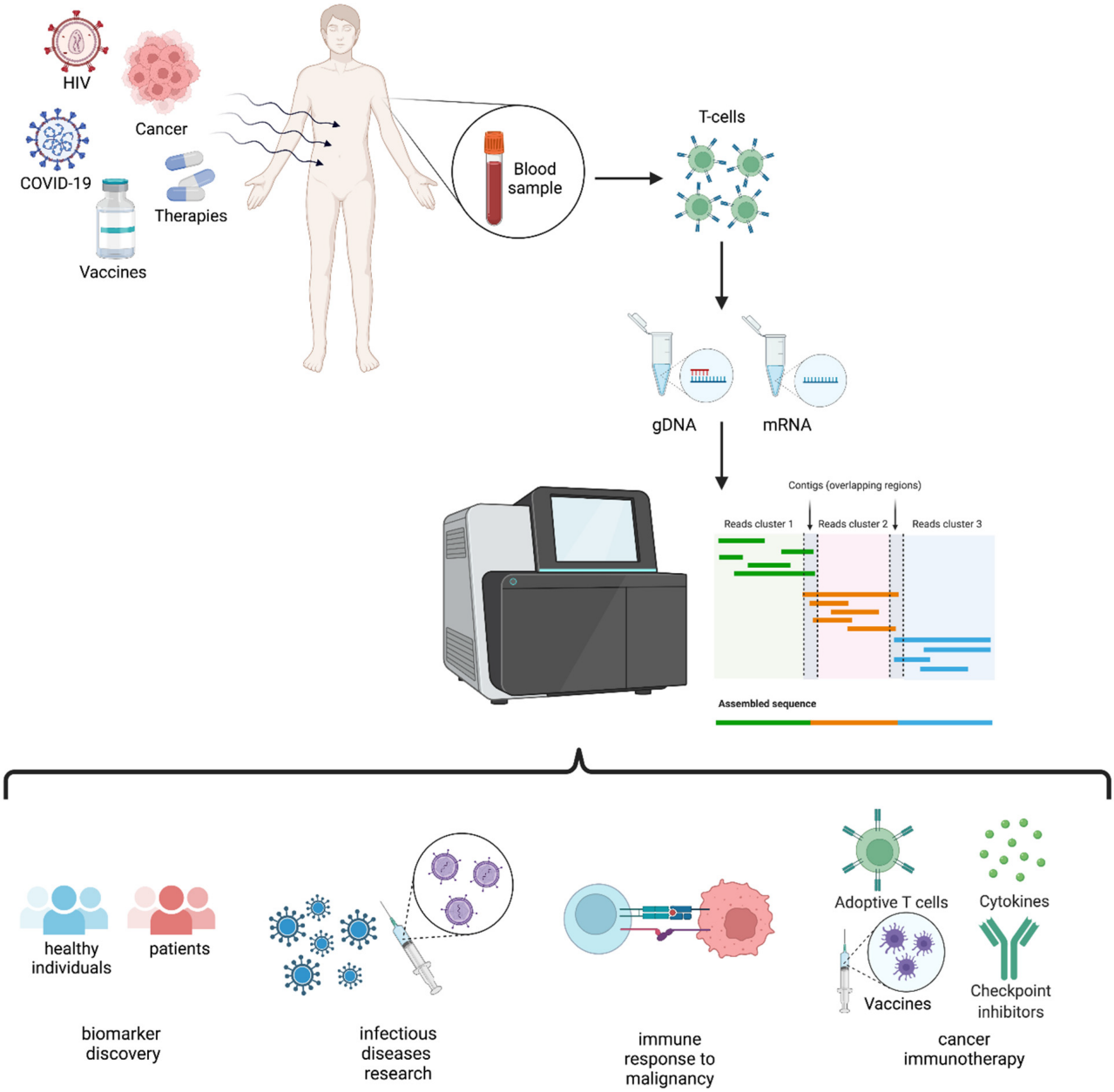

Figure 1. Schematic representation of TCR repertoire analysis and applications.

3. TCR and HLA

As already mentioned, T-cell activation occurs as a consequence of the specific recognition between TCR and foreign antigen peptides presented by the MHC molecules (**Figure 2**), which are transmembrane glycoprotein complexes expressed on the cell surface.

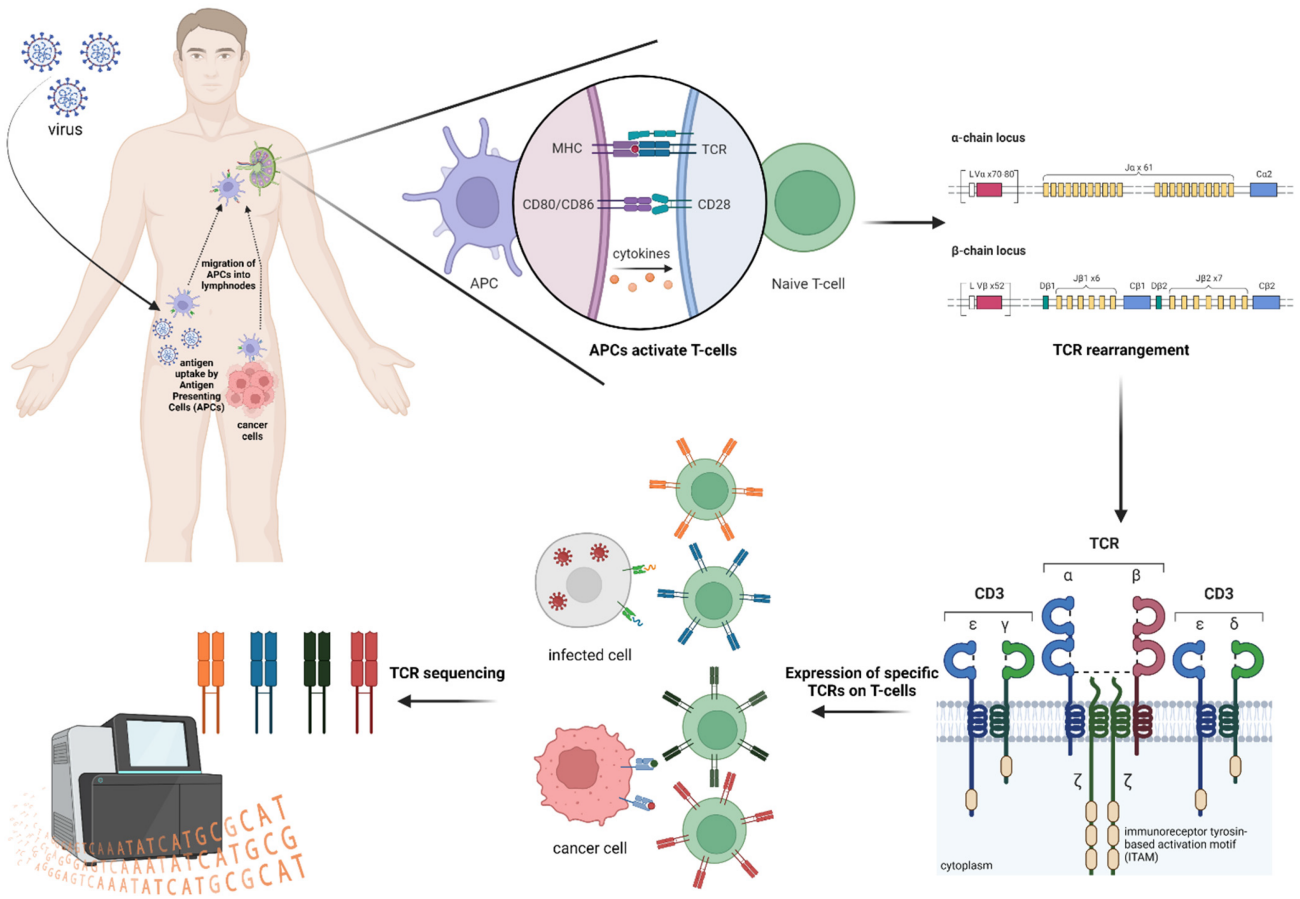

Figure 2. Schematic representation of TCR repertoire generation upon exposure to infectious agents and cancer neoantigens.

MHCs in humans are coded by the highly polymorphic human leukocyte antigen (HLA) gene family located on chromosome 6 and involved in the identification of self versus non-self.

Three subclasses of HLA molecules are expressed in various human tissues: HLA class I, class II and non-classical HLA molecules; some of them make up the class III region.

All nucleated cells express HLA class I proteins on the plasma membrane, allowing to expose peptides derived from intracellular antigens to CD8 T-cells monitoring via TCR interaction. As a result, cells expressing viral or mutated non-self-antigens are killed directly to restrain infection and prevent further cell transformation [31]. There are three main HLA types within this class encoded by the HLA-A, HLA-B and HLA-C loci.

HLA class II proteins are constitutively expressed by professional APCs, including B-cells, and their expression can be upregulated on activated immune cells, binding peptides derived from antigens captured from outside of the cell, presenting ‘exogenous’ peptides to CD4 T-cells [31]. The three main types of HLA in class II are encoded by the HLA-DR, HLA-DQ and HLA-DP loci. APCs can present outside captured antigens also using MHC class I through a process of cross-priming/cross-presentation.

More than 30,000 HLA variants among class I and II have been determined so far, and their permutations raise the number of possible combinations to astronomical numbers, making it unlikely that the individual's resulting HLA type would be shared with an unrelated individual and defining the subset of peptide epitopes that could be presented for immune surveillance [32].

The class III alleles encode for factors involved in the inflammation process, leukocytes differentiation and the complement system [33].

In addition to HLA proteins' role in the human immune system T-cell activation, HLA type plays an important role in driving T-cell positive and negative selection in the thymus, thereby also shaping the naive T-cell repertoire.

HLA importance is evident in the context of organ and bone marrow transplantation, being responsible for the rejection process, but many studies today link the HLA type with disease susceptibility or development and response to therapy for many diseases.

The first link found in this context has been the discovery of HLA-B and Hodgkin lymphoma association [34], and, since then, MHC is considered the genome region with the greatest amount of association with human diseases [35] (some examples are shown in **Table 2**).

Table 2. Examples of association linking HLA type and disease.

Type of HLA Alleles Association	HLA Typing Future Opportunities	Example
With specific infectious diseases or the severity of infection	To provide insight into differences in T-cell repertoires in infectious disease and patterns of T-cell targeting	Heterozygous individuals progress less rapidly to AIDS than HLA homozygous individuals after HIV infection [31]. Kaslow et al. found that HLA-B27 and B57 were strongly associated with slow progression to AIDS [36].
With increased risk of or protection from various autoimmune disorders	To clarify a subject's disease state and potentially stratify patients for treatment studies.	Association of the HLA class I region has been detected for several autoimmune diseases (AIDs); some examples are: - HLA-B with type 1 diabetes (T1D) [37]; - HLA-C with multiple sclerosis (MS) and Graves' disease (GD) [37]; - HLA B-27 with ankylosing spondylitis (AS) [38]; -

Type of HLA Alleles Association	HLA Typing Future Opportunities	Example
		HLA-DRB1, in particular HLA-DRB1*04 and *10 alleles [39] in rheumatoid arthritis (RA); - HLA-G with Crohn's disease (CD) [40].
With cancer therapy outcomes	To understand and infer the efficacy of immunotherapy in specific individuals	Higher heterozygosity in HLA has been linked to a better response to anti-cancer treatments [41].

Rs from a
ease [32].
ssociation
h the fact

Additional analyses revealed that significant TCR clusters, shared within the cohort, may represent markers of immunological memory and showed that most highly HLA-associated TCRs are related to common viral infections, such as influenza virus and Epstein–Barr virus (EBV).

Moreover, they further analyzed CDR3 sequence–HLA allele correlations, identifying a significant negative association between CDR3 and peptide charges, which suggests that the maintenance of charge complementarity across the TCR-MHC complex is a relevant feature of binding.

These results demonstrate the potential of combining statistical tools to TCR repertoires and immune exposure as sequences from the clusters can infer a TCR expansion driver.

Thus, TCR sequence–disease associations are complicated by individual HLA type dependence, thereby characterizing the TCR-HLA interactions. Therefore, it is crucial to understand antigen discrimination by T-cells and to deepen researchers' comprehension of the interplay and associations among individual HLA type, TCR sequences and disease. In this respect, the implications for the development of novel therapeutics are obvious and find translation to many disease settings, including infectious diseases, autoimmune diseases and oncology.

4. TCR Repertoire via HTS: When Details Matter

High-throughput sequencing (HTS) has emerged as a suitable method for evaluating TCR diversity, allowing the characterization of immune repertoires with massive parallel sequencing at a deeper and finer level [8][42]. This technique combines the resolution of individual TCR nucleotide sequences decoded with the ability to read millions of sequences simultaneously [43]. Traditional strategies, such as spectratyping, Sanger sequencing and other assays, such as flow cytometry [30], are time-consuming and insufficient for generating a deep analysis of the immune repertoire.

To perform a TCR repertoire analysis, many aspects must be taken into account, such as the kind of starting material for the library preparation, the method for sequencing [8] and the following data analysis pipeline.

First, following a nucleic acid extraction from the samples cohort of interest, genomic DNA (gDNA) and messenger RNA (mRNA) can both be used for library preparation [44].

The amount of gDNA is proportional to the number of analyzed cells with a 1:1 number of clonotypes and number of cells ratio (1 gDNA template per cell), allowing researchers to determine the relative abundance of sequences in a sample at the cost of unavoidably detecting potentially irrelevant and non-expressed sequences that must be removed through post-processing bioinformatic analysis [45][46].

On the contrary, mRNA is related to cell function/activation [1], and RNA-based methods are more sensitive due to the presence of multiple copies of the transcript of interest per cell. Thus, a more comprehensive recognition of both unique receptor variants and functional expressed TCRs can be obtained using RNA as it allows the detection of very rare clones and reveals sequences effectively transcribed and thus more likely to yield functional TCRs [7][45].

Further, gDNA as input material does not require the reverse transcription step, minimizing the possible biases introduced in cDNA synthesis [5], while starting from already spliced mRNA converted into cDNA holds the advantage that less reverse primers are sufficient for C region amplification, reducing PCR biases from multiplexed J primers [30][42], obtaining both a higher detection sensitivity and no need for adapter sequences [47].

Additionally, RNA-based methods allow the implementation of unique molecular identifiers (UMIs), which consist of random DNA sequences added during cDNA synthesis in order to label individual cDNA molecules, correcting for amplification and sequencing errors [48]. However, since RNA-based approaches are affected by the relative expression of TCRs in the cells and not only by the number of cells expressing the same TCR, those methods are believed to be less reliable in describing the relative abundance of clonotypes in a cell population. The advantages and disadvantages of strategies based on gDNA and mRNA are listed in **Table 3**.

Table 3. DNA-based vs. RNA-based approaches, choosing the right starting material for TCR profiling.

Advantages	
gDNA	**mRNA**
• easier to obtain; • very stable [49]; • no requirement for reverse transcription (RT); • better reflect the number of analyzed cells; • accurate measurement of clonality without bias caused by variable expression levels in different cells.	• higher number of copies in a single cell; • large information at the gene transcription level; • reduced interference of non-coding signals after the splicing process [50];

	• overall length sequence in the CDR region is easily available; • non-productive receptor transcripts are underrepresented [51]. • close proximity of V and C regions after the splicing process facilitates PCR amplification [13].
Disadvantages	
gDNA	**mRNA**
• higher concentration input; • potential annealing of primers for multiple binding sites; • presence of introns and “unused” segments in the sequence of interest that have to be amplified, causing challenges during PCR process [13]; • Detection of all the TCR sequences whether they contribute to a productive or a nonproductive segment arrangement [13].	• introduction of errors during retrotranscription [52]; • easily degraded; • high requirements for extraction, transportation and storage.

uantity of of results required at the end of the process. Currently, peripheral blood is the most used starting material due to the ease and non-invasive sampling procedure, especially in relation to cohorts of healthy subjects, even if peripheral blood lymphocytes are estimated to 2% only of the total lymphocytes in the body [1].

To obtain robust and comparable data, it is important to standardize the processing of all samples to process them as uniformly as possible, starting from a determinate amount of material concentration for each sample and analyzing them with the same parameters as a comparable number of reads in relation to the depth of analysis to achieve in the experiment [53].

The sequencing depth can be adjusted according to the sample type and experimental goals. Deeper sequencing is appropriate when analyzing samples with large or diverse cell populations at the expense of higher throughput.

The majority of TCR repertoire profiling studies are based on the analysis of the CDR3 region; however, full-length sequencing includes additional regions, such as CDR1 and CDR2, involved in antigen receptor binding affinity and/or downstream signaling, and allows to directly clone and express the identified and chosen receptors to perform others experiments. This aspect is crucial when the identification of therapeutic candidate TCRs is the goal of the analysis [45].

TCR HTS methods can be divided into bulk sequencing, for T-cell populations evaluation or single-cell sequencing for the analysis of individual T-cells [2]. The choice between these two analyses depends on the goal of the experiment and on other factors, such as sample requirements, hands-on and total workflow time, degree of polymerase chain reaction (PCR) bias, quantifiability, immune repertoire coverage, ease of data analysis and cost. Generally, for the analysis of immune repertoire diversity in health and disease, a bulk sequencing approach is used because it allows the sampling of many more sequences in a single experiment, even if information about αβ-TCR pairing is lost and undetected low-frequency TCRs could mislead diagnostics outcomes [54].

However, a single-cell approach is preferred in experiments set to investigate the specificity of a TCR for an antigen of interest, for capturing the paired αβ-chains information and producing complete antigen receptors and/or characterizing their function.

Using mRNA, single-cell TCR sequencing makes it possible to evaluate cell transcriptional heterogeneity down to the single nucleotide level and gene expression variability at the single cell level, leveraging the study of phenotypically different cells populations to an unprecedented resolution [55][56]. Compared with TCR bulk sequencing, the number of cells sequenced using the single-cell approach drops to 10^2–10^3 instead of up to 10^6 [54], and another consideration to take into account is that isolating single cells can be challenging, and obtaining viable cells at the end of the process requires care and to work quickly, with a consequent decrease in the number of samples analyzed and an increase in variability in any single-cell study due to the process workflow. If starting from cryopreserved cells, it must be tested whether the process of cryopreservation has changed or damaged the cell viability and/or phenotype. Moreover, data analysis requires specific tools and expertise regarding the most appropriate analytical approaches. Because of the above-mentioned considerations, single-cell is still a more expensive method compared to other sequencing techniques.

Researchers have often performed initial bulk analysis and moved, after selection of features of interest, such as binding affinity, to single-cell [45], even if recently developed commercial single-cell sequencing solutions start to provide full-length paired αβ-chains sequencing of many T-lymphocytes [2]. These ultimate technologies are based, for example, on barcoded gel beads mixed with cells, enzymes and partitioning oil, used to generate V(D)J gene expression libraries (e.g., 10X Genomics, Pleasanton, CA, USA). A major improvement in the throughput came with emulsion-based approaches [7] in which single cells are encapsulated in water-in-oil emulsions, where cDNA is synthetized thanks to TCR primers and RT-PCR reagents and then sequenced [5], maintaining native αβ-chains pairing while sequencing both chains. Quantitative transcriptomics is used to analyze TCRs and other cell markers, and since each cell is individually barcoded, amplification bias is not an issue. Using the same principle of the methods reported above, microfluidic platforms reliant on individual cell compartmentation in microwells or droplets have been applied in single-cell isolation [57].

Droplet-based instruments require a dedicated hardware platform [58], and encapsulation efficiency is variable depending on which method is used, as thousands of cells can be encapsulated, but rare clones can still be missed, while costs remain elevated and hinder a broad application of those approaches [59]. To improve single-cell

TCR sequencing, the cells sorting through the FACS instrument represents an approach to enrich populations of interest through surface markers presence and helps in analyzing rare subsets of T-cells [5].

Library construction and data analysis of bulk and single-cell sequencing approaches share essentially the same principles and workflows [60].

Multiplex PCR represents the most used approach to prepare sequencing libraries for TCR repertoire analysis, and consists of two rounds of PCR using multiple primers, specifically a set of forward primers for V genes and a set of reverse primers for either J or C genes, according to the template used [5].

At first, receptor locus amplification takes place with the addition of known sequences; all the possible recombination events of the receptor sequences are captured using a V and J primers pool in the first PCR, while the additional sequences are fundamental for the incorporation of sequencing adaptors and indexes to each amplicon during the second PCR. By the way, multiple rounds of PCR before sequencing could introduce sequencing biases due to the fact that the priming sequences at the 3′ and 5′ ends of the first PCR overlap significantly between different sites, implying the use of a pool of slightly diverse primers for different TCR sequences amplification, and, importantly, receptor sequences that share similarity with the primers used could be recognized more effectively, affecting clone frequencies' results, with a negative impact on research outcomes [45]. However, it is possible to quantify templates before and after multiplex PCR using synthetic TCR molecules targeted by the multiplexed primers pool, with primer concentration optimization and correction of potential biases [61], and many assays commercially available for library preparation already contain validated internal controls that correct for these biases.

Rapid amplification of cDNA ends is an approach of library preparation used only for RNA templates and relies on a template switching mechanism, an intrinsic property of certain reverse transcriptases (RTs) [62]. This method avoids amplification bias between V regions introduced by multiplex PCR [47], and it is applicable both to 5′ or 3′ ends. The 3′ RACE approach takes advantage of the mRNA poly(A) tail at the 3′ end by using it as a generic priming site for the PCR amplification step following retrotranscription, targeting the region of interest between a known exon and the 3′ end [63].

The 5′ end of mRNA does not present any generic priming sites; therefore, accurate incorporation of an adapter sequence at the cDNA first-strand 5′ end, by adding non-templated nucleotides through RT activity, is required. A hybridization step occurs then between a template-switching oligo complementary to the added non-templated nucleotides, enabling templates to switch and enabling adapter sequence incorporation, which serves the next two PCRs [45]. Consequently, targeting the 5′ adaptor sequence and the C region by using just one pair of primers is enough for all TCR rearrangements' amplification [5], thus reducing PCR errors and ensuring the TCR repertoire profile matches the original sample instead of the primer design. Additionally, a high-on-target rate is guaranteed by the second semi-nested PCR, with a decrease in sequencing costs [45].

Some library prep protocols still employ ligation reaction to anchor adapters and barcodes to the amplicons even if the suboptimal ligation efficiency of the adapters could represent a limiting factor of this choice [62][63], impacting the accuracy of the quantification, especially for the low frequency TCRs, which can justify why 5′ RACE is less reproducible than multiplex PCR [8][64].

After finishing the library preparation step, sequencing of samples can be continued, most of which are run on Illumina platforms. TCR sequences obtained at the end of the workflows consist of sequences of nucleotides that have to be first aligned to VDJ regions' reference sequences and then grouped according to sharing the same CDR3 in order to evaluate clonotypes [65].

The use of algorithms, such as IgBLAST [66], IMGT/HighV-QUEST [67], MiXCR [68], immuneSIM [69] and RTCR [70], allow the evaluation of TCR sequences analogies and discrepancies as compared to publicly available TCR databases.

TCR repertoire analysis is, nowadays, becoming more and more accessible to the scientific community and the pharma industry to unravel TCR specificities, clonality, diversity and the intensity of response associated with treatments and disease states. The panorama of applications for TCR sequencing on the market is really broad and complex, with many companies proposing specific protocols according to the specificities previously defined. As a means of orienting researchers in the choice of the best suited approach for TCR sequencing in the present scenario researchers provide, to the best of researchers' knowledge, a comprehensive and synthetic description of the kits and the services accessible on the market at the moment.

References

1. Liu, H.; Pan, W.; Tang, C.; Tang, Y.; Wu, H.; Yoshimura, A.; Deng, Y.; He, N.; Li, S. The Methods and Advances of Adaptive Immune Receptors Repertoire Sequencing. Theranostics 2021, 11, 8945–8963.
2. Dahal-Koirala, S.; Balaban, G.; Neumann, R.S.; Scheffer, L.; Lundin, K.E.A.; Greiff, V.; Sollid, L.M.; Qiao, S.-W.; Sandve, G.K. TCRpower: Quantifying the Detection Power of T-Cell Receptor Sequencing with a Novel Computational Pipeline Calibrated by Spike-in Sequences. Brief. Bioinform. 2022, 23, bbab566.
3. Li, N.; Yuan, J.; Tian, W.; Meng, L.; Liu, Y. T-cell Receptor Repertoire Analysis for the Diagnosis and Treatment of Solid Tumor: A Methodology and Clinical Applications. Cancer Commun. 2020, 40, 473–483.
4. Aversa, I.; Malanga, D.; Fiume, G.; Palmieri, C. Molecular T-Cell Repertoire Analysis as Source of Prognostic and Predictive Biomarkers for Checkpoint Blockade Immunotherapy. Int. J. Mol. Sci. 2020, 21, 2378.

5. Pai, J.A.; Satpathy, A.T. High-Throughput and Single-Cell T Cell Receptor Sequencing Technologies. Nat. Methods 2021, 18, 881–892.
6. Watkins, T.S.; Miles, J.J. The Human T-cell Receptor Repertoire in Health and Disease and Potential for Omics Integration. Immunol. Cell Biol. 2021, 99, 135–145.
7. Chiffelle, J.; Genolet, R.; Perez, M.A.; Coukos, G.; Zoete, V.; Harari, A. T-Cell Repertoire Analysis and Metrics of Diversity and Clonality. Curr. Opin. Biotechnol. 2020, 65, 284–295.
8. Rosati, E.; Dowds, C.M.; Liaskou, E.; Henriksen, E.K.K.; Karlsen, T.H.; Franke, A. Overview of Methodologies for T-Cell Receptor Repertoire Analysis. BMC Biotechnol. 2017, 17, 61.
9. Yamauchi, M. Mechanisms Underlying the Suppression of Chromosome Rearrangements by Ataxia-Telangiectasia Mutated. Genes 2021, 12, 1232.
10. Nishana, M.; Raghavan, S.C. Role of Recombination Activating Genes in the Generation of Antigen Receptor Diversity and Beyond. Immunology 2012, 137, 271–281.
11. Yang, G.; Ou, M.; Chen, H.; Guo, C.; Chen, J.; Lin, H.; Tang, D.; Xue, W.; Li, W.; Sui, W.; et al. Characteristic Analysis of TCR β-Chain CDR3 Repertoire for Pre- and Post-Liver Transplantation. Oncotarget 2018, 9, 34506–34519.
12. Bassing, C.H.; Swat, W.; Alt, F.W. The Mechanism and Regulation of Chromosomal V(D)J Recombination. Cell 2002, 109, S45–S55.
13. Calis, J.J.A.; Rosenberg, B.R. Characterizing Immune Repertoires by High Throughput Sequencing: Strategies and Applications. Trends Immunol. 2014, 35, 581–590.
14. Valkiers, S.; de Vrij, N.; Gielis, S.; Verbandt, S.; Ogunjimi, B.; Laukens, K.; Meysman, P. Recent Advances in T-Cell Receptor Repertoire Analysis: Bridging the Gap with Multimodal Single-Cell RNA Sequencing. ImmunoInformatics 2022, 5, 100009.
15. Pauken, K.E.; Lagattuta, K.A.; Lu, B.Y.; Lucca, L.E.; Daud, A.I.; Hafler, D.A.; Kluger, H.M.; Raychaudhuri, S.; Sharpe, A.H. TCR-Sequencing in Cancer and Autoimmunity: Barcodes and Beyond. Trends Immunol. 2022, 43, 180–194.
16. Robins, H.S.; Campregher, P.V.; Srivastava, S.K.; Wacher, A.; Turtle, C.J.; Kahsai, O.; Riddel, S.R.; Warren, E.H.; Carlson, C.S. Comprehensive Assessment of T-Cell Receptor β-Chain Diversity in Aβ T Cells. Blood 2009, 114, 4099–4107.
17. Warren, R.L.; Freeman, J.D.; Zeng, T.; Choe, G.; Munro, S.; Moore, R.; Webb, J.R.; Holt, R.A. Exhaustive T-Cell Repertoire Sequencing of Human Peripheral Blood Samples Reveals Signatures of Antigen Selection and a Directly Measured Repertoire Size of at Least 1 Million Clonotypes. Genome Res. 2011, 21, 790–797.
18. Nikolich-Žugich, J.; Slifka, M.K.; Messaoudi, I. The Many Important Facets of T-Cell Repertoire Diversity. Nat. Rev. Immunol. 2004, 4, 123–132.

19. Dupic, T.; Marcou, Q.; Walczak, A.M.; Mora, T. Genesis of the Aβ T-Cell Receptor. PLoS Comput. Biol. 2019, 15, e1006874.
20. Sewell, A.K. Why Must T Cells Be Cross-Reactive? Nat. Rev. Immunol. 2012, 12, 669–677.
21. Kumar, B.V.; Connors, T.J.; Farber, D.L. Human T Cell Development, Localization, and Function throughout Life. Immunity 2018, 48, 202–213.
22. Gutierrez, L.; Beckford, J.; Alachkar, H. Deciphering the TCR Repertoire to Solve the COVID-19 Mystery. Trends Pharmacol. Sci. 2020, 41, 518–530.
23. Chang, J.T.; Wherry, E.J.; Goldrath, A.W. Molecular Regulation of Effector and Memory T Cell Differentiation. Nat. Immunol. 2014, 15, 1104–1115.
24. Daniels, M.A.; Teixeiro, E. TCR Signaling in T Cell Memory. Front. Immunol. 2015, 6, 617.
25. Zvyagin, I.V.; Pogorelyy, M.V.; Ivanova, M.E.; Komech, E.A.; Shugay, M.; Bolotin, D.A.; Shelenkov, A.A.; Kurnosov, A.A.; Staroverov, D.B.; Chudakov, D.M.; et al. Distinctive Properties of Identical Twins' TCR Repertoires Revealed by High-Throughput Sequencing. Proc. Natl. Acad. Sci. USA 2014, 111, 5980–5985.
26. Chen, G.; Yang, X.; Ko, A.; Sun, X.; Gao, M.; Zhang, Y.; Shi, A.; Mariuzza, R.A.; Weng, N. Sequence and Structural Analyses Reveal Distinct and Highly Diverse Human CD8 + TCR Repertoires to Immunodominant Viral Antigens. Cell Rep. 2017, 19, 569–583.
27. Pantaleo, G.; Demarest, J.F.; Soudeyns, H.; Graziosi, C.; Denis, F.; Adelsberger, J.W.; Borrow, P.; Saag, M.S.; Shaw, G.M.; Sekaly, R.P.; et al. Major Expansion of CD8+ T Cells with a Predominant vp Usage during the Primary Immune Response to HIV. Nature 1994, 370, 463–467.
28. Li, B.; Li, T.; Pignon, J.-C.; Wang, B.; Wang, J.; Shukla, S.A.; Dou, R.; Chen, Q.; Hodi, F.S.; Choueiri, T.K.; et al. Landscape of Tumor-Infiltrating T Cell Repertoire of Human Cancers. Nat. Genet. 2016, 48, 725–732.
29. Holtmeier, W.; Kabelitz, D.; Gammadelta, T. Cells Link Innate and Adaptive Immune Responses. Chem. Immunol. Allergy 2005, 86, 151–183.
30. Woodsworth, D.J.; Castellarin, M.; Holt, R.A. Sequence Analysis of T-Cell Repertoires in Health and Disease. Genome Med. 2013, 5, 98.
31. Mosaad, Y.M. Clinical Role of Human Leukocyte Antigen in Health and Disease. Scand. J. Immunol. 2015, 82, 283–306.
32. DeWitt, W.S.; Smith, A.; Schoch, G.; Hansen, J.A.; Matsen, F.A.; Bradley, P. Human T Cell Receptor Occurrence Patterns Encode Immune History, Genetic Background, and Receptor Specificity. eLife 2018, 7, e38358.

33. Dendrou, C.A.; Petersen, J.; Rossjohn, J.; Fugger, L. HLA Variation and Disease. Nat. Rev. Immunol. 2018, 18, 325–339.
34. Amiel, J. Study of the Leukocyte Phenotypes in Hodgkin's Disease in Histocompatibility Testing; Teraski, P.I., Ed.; Munksgaard: Copenhagen, Denmark, 1967; pp. 79–81.
35. Trowsdale, J.; Knight, J.C. Major Histocompatibility Complex Genomics and Human Disease. Annu. Rev. Genom. Hum. Genet. 2013, 14, 301–323.
36. Kaslow, R.A.; Carrington, M.; Apple, R.; Park, L.; Muñoz, A.; Saah, A.J.; Goedert, J.J.; Winkler, C.; O'Brien, S.J.; Rinaldo, C.; et al. Influence of Combinations of Human Major Histocompatibility Complex Genes on the Course of HIV-1 Infection. Nat. Med. 1996, 2, 405–411.
37. Simmonds, M.; Gough, S. The HLA Region and Autoimmune Disease: Associations and Mechanisms of Action. Curr. Genom. 2007, 8, 453–465.
38. Flores-Robles, B.-J.; Labrador-Sánchez, E.; Andrés-Trasahedo, E.; Pinillos-Aransay, V.; Joven-Zapata, M.-Y.; Torrecilla Lerena, L.; Salazar-Asencio, O.-A.; López-Martín, J.-A. Concurrence of Rheumatoid Arthritis and Ankylosing Spondylitis: Analysis of Seven Cases and Literature Review. Case Rep. Rheumatol. 2022, 2022, 8500567.
39. Padyukov, L. Genetics of Rheumatoid Arthritis. Semin. Immunopathol. 2022, 44, 47–62.
40. Abdul-Hussein, S.S.; Ali, E.N.; Zaki, N.H.; Ad'hiah, A.H. Genetic Polymorphism of HLA-G Gene (G*01:03, G*01:04, and G*01:05N) in Iraqi Patients with Inflammatory Bowel Disease (Ulcerative Colitis and Crohn's Disease). Egypt J. Med. Hum. Genet. 2021, 22, 34.
41. Chowell, D.; Morris, L.G.T.; Grigg, C.M.; Weber, J.K.; Samstein, R.M.; Makarov, V.; Kuo, F.; Kendall, S.M.; Requena, D.; Riaz, N.; et al. Patient HLA Class I Genotype Influences Cancer Response to Checkpoint Blockade Immunotherapy. Science 2018, 359, 582–587.
42. De Simone, M.; Rossetti, G.; Pagani, M. Single Cell T Cell Receptor Sequencing: Techniques and Future Challenges. Front. Immunol. 2018, 9, 1638.
43. Metzker, M.L. Sequencing Technologies—The next Generation. Nat. Rev. Genet. 2010, 11, 31–46.
44. Rizzo, J.M.; Buck, M.J. Key Principles and Clinical Applications of "Next-Generation" DNA Sequencing. Cancer Prev. Res. 2012, 5, 887–900.
45. Takara Bio Blog Team. 4 Factors to Consider for Immune Repertoire Profiling; Web Document Reprint; Takara Bio USA, Inc.: San Jose, CA, USA, 2019.
46. Kockelbergh, H.; Evans, S.; Deng, T.; Clyne, E.; Kyriakidou, A.; Economou, A.; Luu Hoang, K.N.; Woodmansey, S.; Foers, A.; Fowler, A.; et al. Utility of Bulk T-Cell Receptor Repertoire Sequencing Analysis in Understanding Immune Responses to COVID-19. Diagnostics 2022, 12, 1222.

47. Barennes, P.; Quiniou, V.; Shugay, M.; Egorov, E.S.; Davydov, A.N.; Chudakov, D.M.; Uddin, I.; Ismail, M.; Oakes, T.; Chain, B.; et al. Benchmarking of T Cell Receptor Repertoire Profiling Methods Reveals Large Systematic Biases. Nat. Biotechnol. 2021, 39, 236–245.
48. Shugay, M.; Britanova, O.V.; Merzlyak, E.M.; Turchaninova, M.A.; Mamedov, I.Z.; Tuganbaev, T.R.; Bolotin, D.A.; Staroverov, D.B.; Putintseva, E.K.; Plevova, K.; et al. Towards Error-Free Profiling of Immune Repertoires. Nat. Methods 2014, 11, 653–655.
49. Logan, A.C.; Gao, H.; Wang, C.; Sahaf, B.; Jones, C.D.; Marshall, E.L.; Buno, I.; Armstrong, R.; Fire, A.Z.; Weinberg, K.I.; et al. High-Throughput VDJ Sequencing for Quantification of Minimal Residual Disease in Chronic Lymphocytic Leukemia and Immune Reconstitution Assessment. Proc. Natl. Acad. Sci. USA 2011, 108, 21194–21199.
50. Tiller, T.; Busse, C.E.; Wardemann, H. Cloning and Expression of Murine Ig Genes from Single B Cells. J. Immunol. Methods 2009, 350, 183–193.
51. Li, S.; Wilkinson, M.F. Nonsense Surveillance in Lymphocytes? Immunity 1998, 8, 135–141.
52. Wang, C.; Liu, Y.; Xu, L.T.; Jackson, K.J.L.; Roskin, K.M.; Pham, T.D.; Laserson, J.; Marshall, E.L.; Seo, K.; Lee, J.; et al. Effects of Aging, Cytomegalovirus Infection, and EBV Infection on Human B Cell Repertoires. J. Immunol. 2014, 192, 603–611.
53. Mamedov, I.Z.; Britanova, O.V.; Zvyagin, I.V.; Turchaninova, M.A.; Bolotin, D.A.; Putintseva, E.V.; Lebedev, Y.B.; Chudakov, D.M. Preparing Unbiased T-Cell Receptor and Antibody CDNA Libraries for the Deep Next Generation Sequencing Profiling. Front. Immunol. 2013, 4, 456.
54. Trück, J.; Eugster, A.; Barennes, P.; Tipton, C.M.; Luning Prak, E.T.; Bagnara, D.; Soto, C.; Sherkow, J.S.; Payne, A.S.; Lefran, M.; et al. Biological Controls for Standardization and Interpretation of Adaptive Immune Receptor Repertoire Profiling. eLife 2021, 10, e66274.
55. Yip, S.H.; Sham, P.C.; Wang, J. Evaluation of Tools for Highly Variable Gene Discovery from Single-Cell RNA-Seq Data. Brief. Bioinform. 2019, 20, 1583–1589.
56. Olsen, T.K.; Baryawno, N. Introduction to Single-Cell RNA Sequencing. Curr. Protoc. Mol. Biol. 2018, 122, e57.
57. Prakadan, S.M.; Shalek, A.K.; Weitz, D.A. Scaling by Shrinking: Empowering Single-Cell “omics” with Microfluidic Devices. Nat. Rev. Genet. 2017, 18, 345–361.
58. Chen, G.; Ning, B.; Shi, T. Single-Cell RNA-Seq Technologies and Related Computational Data Analysis. Front. Genet. 2019, 10, 317.
59. Salomon, R.; Kaczorowski, D.; Valdes-Mora, F.; Nordon, R.E.; Neild, A.; Farbehi, N.; Bartonicekcg, N.; Gallego-Ortega, D. Droplet-Based Single Cell RNAseq Tools: A Practical Guide. Lab Chip 2019, 19, 1706–1727.

60. Zhang, J.; Song, C.; Tian, Y.; Yang, X. Single-Cell RNA Sequencing in Lung Cancer: Revealing Phenotype Shaping of Stromal Cells in the Microenvironment. Front. Immunol. 2022, 12, 802080.
61. Carlson, C.S.; Emerson, R.O.; Sherwood, A.M.; Desmarais, C.; Chung, M.-W.; Parsons, J.M.; Steen, M.S.; LaMadrid-Herrmannsfeldt, M.A.; Williamson, D.W.; Livingston, R.J.; et al. Using Synthetic Templates to Design an Unbiased Multiplex PCR Assay. Nat. Commun. 2013, 4, 2680.
62. Wulf, M.G.; Maguire, S.; Humbert, P.; Dai, N.; Bei, Y.; Nichols, N.M.; Correa, I.R., Jr.; Guan, S. Non-Templated Addition and Template Switching by Moloney Murine Leukemia Virus (MMLV)-Based Reverse Transcriptases Co-Occur and Compete with Each Other. J. Biol. Chem. 2019, 294, 18220–18231.
63. Adamopoulos, P.G.; Tsiakanikas, P.; Stolidi, I.; Scorilas, A. A Versatile 5′ RACE-Seq Methodology for the Accurate Identification of the 5′ Termini of MRNAs. BMC Genom. 2022, 23, 163.
64. Liu, X.; Zhang, W.; Zeng, X.; Zhang, R.; Du, Y.; Hong, X.; Cao, H.; Su, Z.; Wang, C.; Wu, J.; et al. Systematic Comparative Evaluation of Methods for Investigating the TCRβ Repertoire. PLoS ONE 2016, 11, e0152464.
65. Aoki, H.; Shichino, S.; Matsushima, K.; Ueha, S. Revealing Clonal Responses of Tumor-Reactive T-Cells Through T Cell Receptor Repertoire Analysis. Front. Immunol. 2022, 13, 807696.
66. Ye, J.; Ma, N.; Madden, T.L.; Ostell, J.M. IgBLAST: An Immunoglobulin Variable Domain Sequence Analysis Tool. Nucleic Acids Res. 2013, 41, W34–W40.
67. Alamyar, E.; Duroux, P.; Lefranc, M.P.; Giudicelli, V. IMGT ® Tools for the Nucleotide Analysis of Immunoglobulin (IG) and T Cell Receptor (TR) V- (D)-J Repertoires, Polymorphisms, and IG Mutations: IMGT/V-QUEST and IMGT/HighV-QUEST for NGS. Methods Mol. Biol. 2012, 882, 569–604.
68. Bolotin, D.A.; Poslavsky, S.; Mitrophanov, I.; Shugay, M.; Mamedov, I.Z.; Putintseva, E.V.; Chudakov, D.M. MiXCR: Software for Comprehensive Adaptive Immunity Profiling. Nat. Methods 2015, 12, 380–381.
69. Weber, C.R.; Akbar, R.; Yermanos, A.; Pavlović, M.; Snapkov, I.; Sandve, G.K.; Reddy, S.T.; Greiff, V. ImmuneSIM: Tunable Multi-Feature Simulation of B- and T-Cell Receptor Repertoires for Immunoinformatics Benchmarking. Bioinformatics 2020, 36, 3594–3596.
70. Gerritsen, B.; Pandit, A.; Andeweg, A.C.; de Boer, R.J. RTCR: A Pipeline for Complete and Accurate Recovery of T Cell Repertoires from High Throughput Sequencing Data. Bioinformatics 2016, 32, 3098–3106.