DISCOVER PATTERNS IMPRINTED
IN OUR IMMUNE SYSTEM

Immune sequencing with iRepertoire

SINGLE CELL

BULK REPERTOIRE

BIOINFORMATICS

AUTOMATION

iRepertoire, Inc.

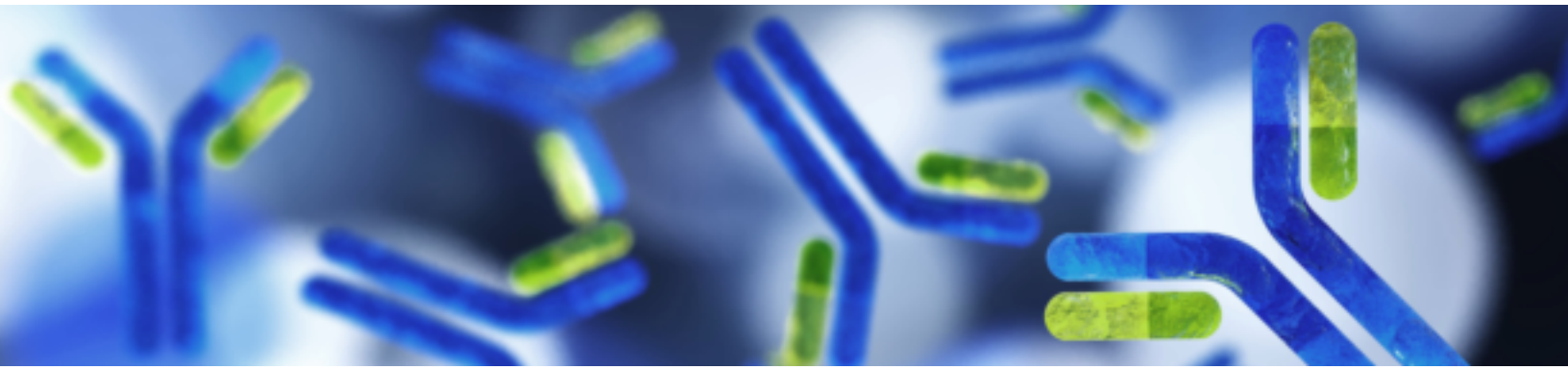
CONTENTS

Explore more content at iRepertoire.com

CHAPTER ONE

How to sequence the immune repertoire

REVEAL IMPORTANT INFORMATION ABOUT THE PATHOLOGY AND PROGNOSIS OF DISEASE



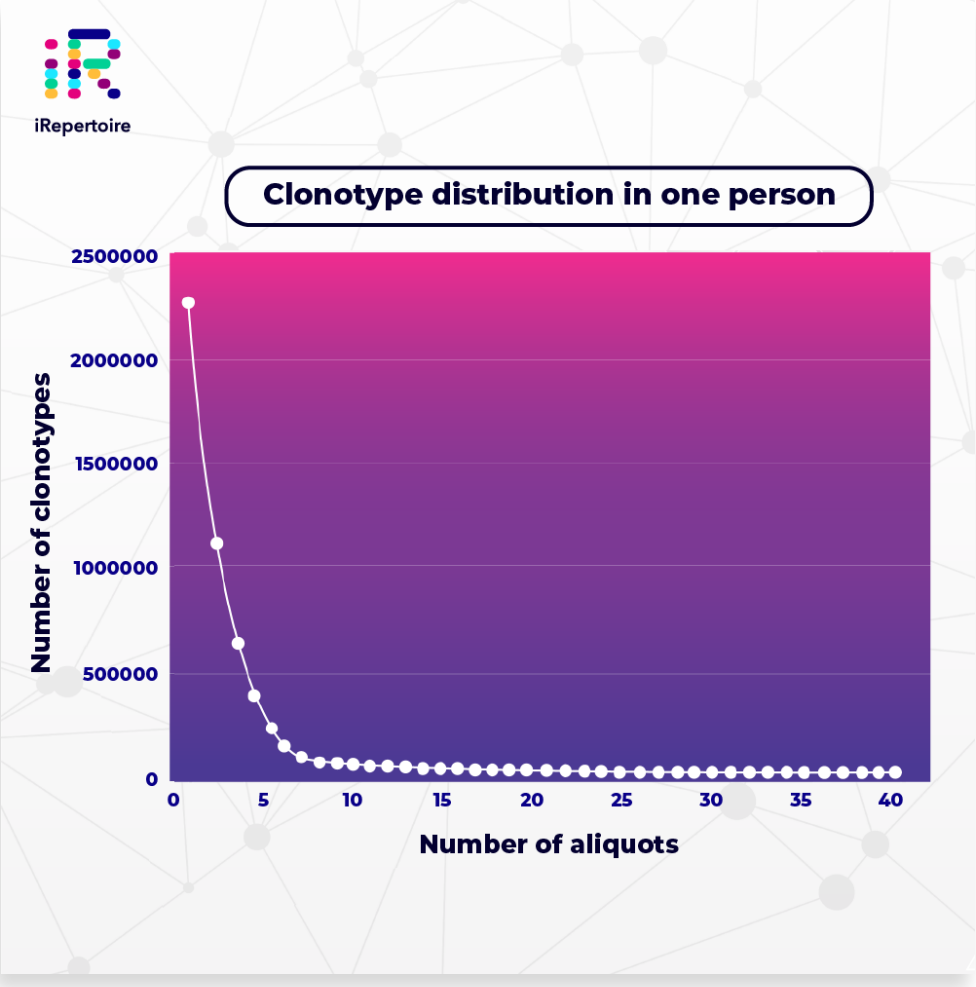
Immune repertoire sequencing requires a precise approach

Capturing the **full clonal diversity of the immune repertoire** can reveal important information about the pathology and prognosis of disease. Compared to normal individuals, the immune repertoire of patients with different diseases may be quantitatively or qualitatively different from healthy controls in terms of composition and diversity. Qualitative changes may present as increased sharing of disease-specific CDR3s in T or B cells. Quantitative changes may manifest as increases and decreases in repertoire diversity. Capturing both qualitative and quantitative changes requires a **repertoire sequencing approach** that is both sensitive (so as to include all CDR3s) and unbiased (to allow for relative quantitation of all CDR3s).

For examples of how immune repertoire sequencing applies to health, see the Learning Cener Article **“Why sequence the immune repertoire?”** also included in **eBook 1: An intro to immunology and immune sequencing**.

Quantitative and inclusive immune repertoire analysis

While there are hundreds of thousands of different clonotypes in one person's immune system, only a small number of clonotypes make up the bulk of the immune repertoire. Sampling the CDR3 clonotypes in a single person would thus produce a distribution like the one shown in the graph below:



This long-tailed distribution results in a low signal-to-noise ratio, as the more frequent CDR3s will be overrepresented in a sequencing library. To effectively capture the full breadth and depth of the immune repertoire, the sequencing methods must, therefore, be inclusive and quantitative.

Information about the absolute quantity of particular clonotypes can be difficult to attain because amplification is required to prepare libraries for sequencing. The amplification process is biased towards abundant and small amplicons, which can cause less frequent, rare amplicons to be underrepresented in the final library. iRepertoire has developed two different multiplexed PCR methods to reduce these biases.

These technologies are described in **Chapter Two on iRepertoire's amplification technologies.**

Deciding what to sequence

Depending on the goals of the study, an experimenter wishing to sequence an immune repertoire will need to determine what region(s) of the T cell and/or B cell receptor (TCR and BCR) genes to target, whether to use DNA or RNA as starting material, and whether to sequence bulk tissue or **single cells.**

The regions of TCRs and BCRs that play the most important role in antigen recognition are also the most variable. These variable regions are constructed from a random shuffling of gene segments, and, in the case of BCRs, somatic mutations. The most variable portion of the variable region is the CDR3 region. It is, therefore, critically important to target the CDR3 region in immune repertoire sequencing. All of **iRepertoire's primer systems** cover the highly variable CDR3 region.

Sequencing the full variable region can provide more information, but that extra information comes at the cost of increased sequencing reads or decreased sequencing coverage. For more about sequencing coverage, see our post about **choosing read depth.**

While both RNA and DNA sequencing can provide a picture of the immune repertoire, only RNA sequencing will reflect the library of TCR and BCR sequences that are actually expressed (for more information, see our post on **template selection**).

Finally, bulk sequencing is required to analyze immune repertoire diversity, but information about the cognate pairing of BCR/TCR chains within a given B or T cell is lost in bulk sequencing. Researchers that wish to examine the relationships between paired chains, in particular clonotypes, should perform single cell sequencing. For more information about how to sequence single cells, see our post on **cell sorting**.

CHAPTER TWO

iRepertoire's amplification technologies

FULLY CAPTURE THE DIVERSITY OF THE IMMUNE REPERTOIRE

Challenges of immune repertoire sequencing

The unique challenges involved with sequencing the immune repertoire have necessitated the development of repertoire-specific library preparation technologies. In order to fully capture the diversity of the immune repertoire, amplification technologies have to be used that account for biases towards small amplicons and amplifying rare clonotypes beyond the templates that are naturally high in abundance. This is no small feat considering there are multiple chain types that contribute to each individual clonotype in a given repertoire, and the diversity of the repertoire has to account for gene recombination, expression, and hypermutation.

iRepertoire's multiplex PCR amplification technologies were specially designed to amplify the B cell receptor (BCR) and T cell receptor (TCR) chains of the immune system with both inclusivity and quantitation in mind.

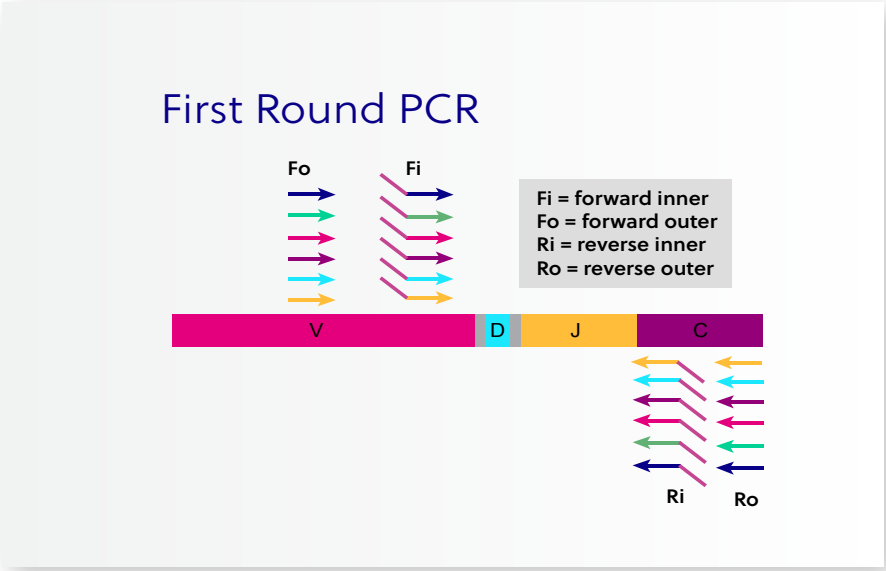
iRepertoire's multiplexing solution

iRepertoire offers two different **multiplex PCR approaches**, both of which amplify all of the V(D)Js in a sample, including the highly variable CDR3 region. With arm-PCR (amplicon rescued multiplex PCR, patent 7,999,092) you can amplify the chain of your choosing from human or mouse samples with superior sensitivity so as to include all the diversity present in a sample, even for extremely rare clonotypes. With dam-PCR (dimer avoided multiplex PCR, patent pending), you can amplify any or all of the BCR and TCR chains in a much more quantitative manner than competing approaches.

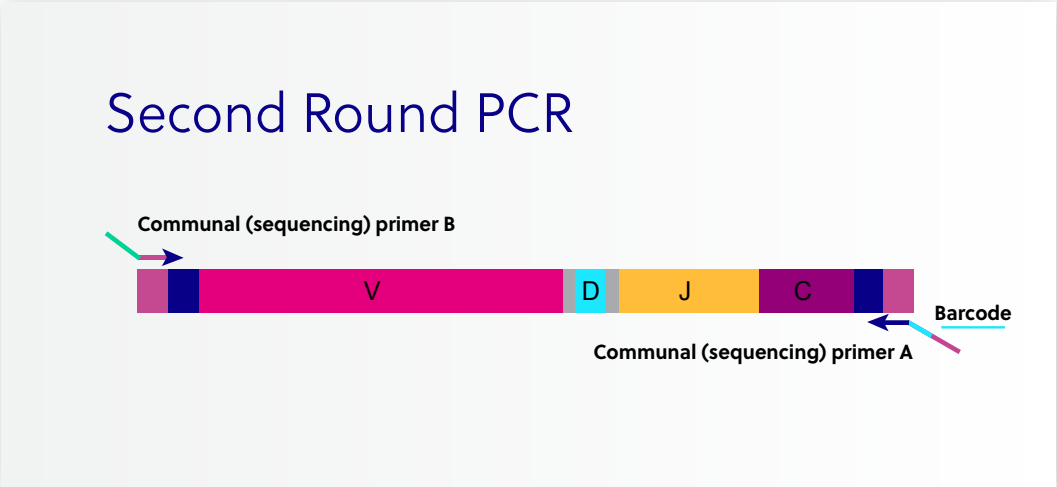
Amplicon rescued multiplex PCR (arm-PCR)

Arm-PCR uses a high concentration of **hundreds of nested inside and outside gene-specific primers** in the initial combined reverse transcription (RT) plus PCR round one. Both the reverse outside and inside primers can contribute to first strand synthesis during RT, which is especially important if there are any secondary RNA structures that make the inner primer binding site inaccessible during RT.

The outside primers help to improve the sensitivity of the reaction by increasing target template abundance for the inside primer to bind. Because the nested primer mix goes through many binding and extension cycles, **arm-PCR is a great technological solution for rare clonotype discovery**. When RT-PCR1 is complete, target amplicons are rescued, and PCR round two is performed using fresh enzymes. For PCR2, communal primers that recognize the shared tag sequence introduced during the first round of amplification are used for further amplification.



In the first round of arm-PCR, inside and outside TCR/BCR-specific primers amplify the targets of interest, increasing sensitivity and appending communal primer binding sites.



In the second round of arm-PCR, communal primers exponentially amplify the target amplicons generated in the first round

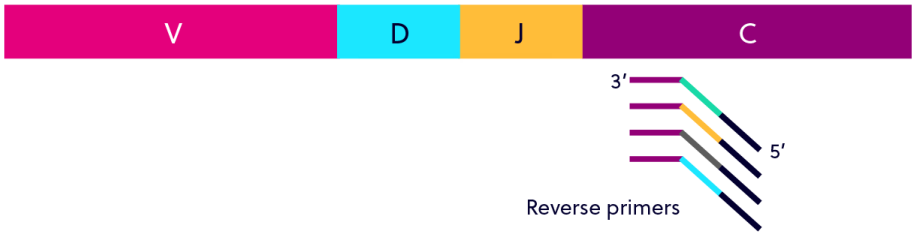
Learn more about arm-PCR

Dimer avoided multiplex PCR (dam-PCR)

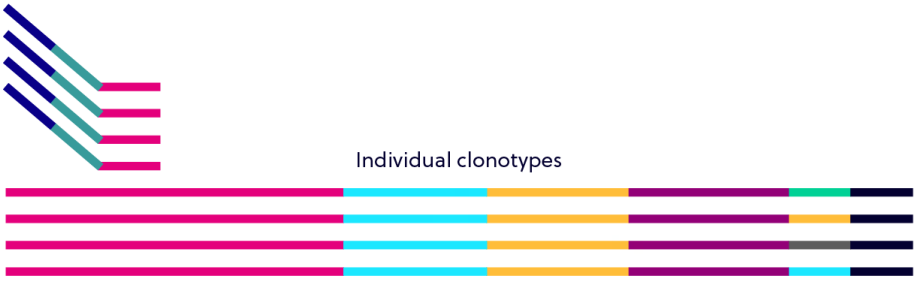
iRepertoire's dam-PCR technology lets you select any combination of TCR and BCR chains for simultaneous amplification in one reaction. This is made possible by the unique single-cycle binding and extension steps and stringent clean up steps in-between, which omit harmful dimer formation. First, only the 3' primer is added and one binding and extension step is performed. The unincorporated 3' primers are washed away, and then the 5' primers are added. After another single cycle binding and extension protocol, the 5' primers are washed away. Primers that bind to the communal primer sites introduced in the first two steps are added for multi-cycle, exponential amplification.

Dam-PCR also allows for the inclusion of unique molecular indices (UMI) so that each strand of RNA can be tagged for direct quantification and both PCR and/or sequencing error removal. This increases confidence in the sequenced targets, and the quantification allows for investigation of interchain ratios in the immune adaptome. Thus, while arm-PCR provides an inclusive, semi-quantitative overview of the immune repertoire with respect to particular chains, dam-PCR provides a more quantitative look at the frequency of particular clonotypes of interest for any or all BCR and TCR chains.

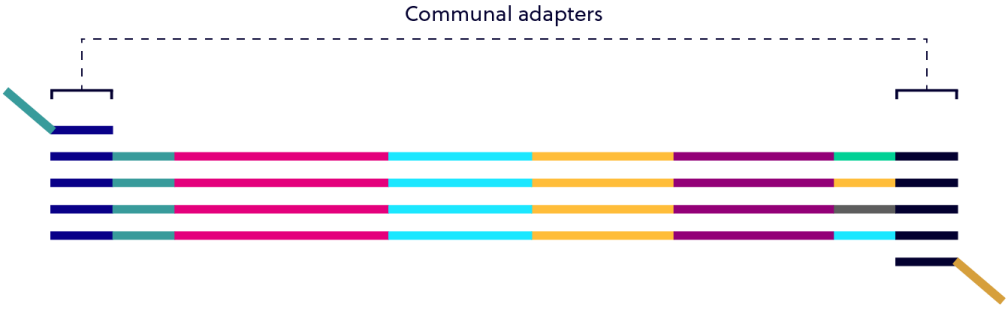
Reverse transcription and first strand synthesis



Second strand synthesis



Two rounds of enrichment



Dimer-avoided multiplex (dam) PCR removes unwanted side products, including primer dimers, and allows for greater sensitivity and efficiency in amplifying nucleic acids. dam-PCR comprises three steps: first, a gene specific reverse primer mix is used for reverse transcription and first strand cDNA synthesis in a single cycle binding and extension step. Second, at the end of the cycle, unused residual primers are removed. In the second step, a gene specific forward primer mix performs second strand cDNA synthesis in another single cycle binding and extension step. All residual unused primers are again removed. In the third step, the communal adapter sequence from steps one and two are used for two further rounds of amplification and enrichment. Given the selection of DNA strands and removal of unused primers prior to amplification, dam-PCR avoids primer-dimer formation, helps reduce bias, and avoids off-target amplicons while increasing the signal of specific targets.

Automation with iR-Flex

iRepertoire has also developed an automation platform called iR-Flex to enable immune repertoire amplification by arm-PCR or dam-PCR in enclosed, contamination-free cassettes. Because dam-PCR requires extensive sample processing, dam-PCR is only offered via iR-Flex or through **services**.

Learn more about iR-Flex

iR-Cassette and iR-Flex Automation System

CHAPTER THREE

iRepertoire's primer systems

OPTIMIZED PRIMER SYSTEMS FOR EVERY V(D)J TARGET

Overview

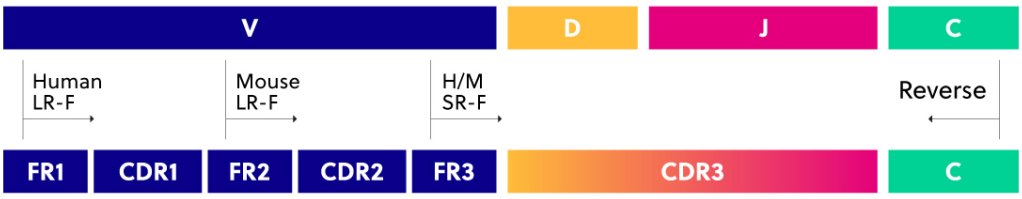
iRepertoire developed multiplex PCR primer mixes for V(D)J amplification and sequencing for both our arm-PCR and dam-PCR technologies. In these multiplex PCR systems, a nested set of inside and outside primers has been designed and optimized for **every V(D)J target** needing to be co-amplified in the multiplex assay. We offer primers, amplification services, and **sequencing services for all TCR and BCR chains** for both human and mouse species in order to meet your research needs, whatever they may be.

For more information about arm-PCR and dam-PCR, see the **previous chapter on iRepertoire's amplification technologies.**

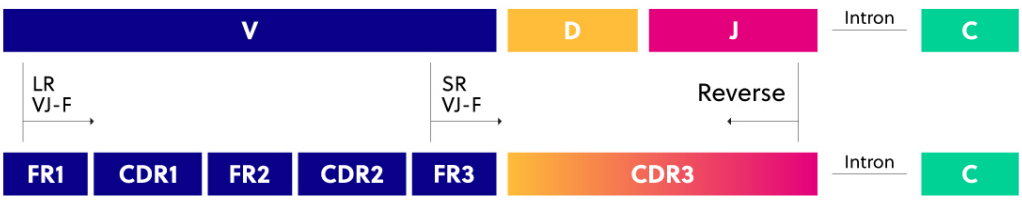
iRepertoire offers several different primer systems

iRepertoire has multiple different primer systems that vary by the regions targeted, the desired read length, and the species. Regardless of what primer system works best for your study, all of our primers cover the highly variable CDR3 region at a minimum, as this is generally the greatest area of diversity and interest.

RNA Primer Systems



gDNA Primer System - Human TCR-Beta only



Our primer technology essentially grew alongside sequencing and read depth ability. Our services began with our short-read (SR) human and mouse primers, which cover from within Framework-3 (FR3) into the Constant Region (C). These SR primers are compatible with 100/150 paired end read (PER) sequencing. As sequencing technology became more advanced, our long-read (LR) primer systems became available. These cover from within FR1 for human samples and FR2 for mouse, and both continue through to the C region. iRepertoire's LR primers are compatible with 250 PER sequencing.

While we typically prefer RNA as the starting template (see our Learning Center article on **choice of starting template**), we understand some researcher's preference for gDNA. Therefore, we have our V-J primer system for use with gDNA as the starting template (available for Human TCR-Beta only).

The original V-J primer system (SR-VJ) covers from within FR3 to the end of the J-gene and is compatible with 100/150 PER sequencing. A new long read V-J primer system (LR-VJ) is currently under development and should be released by Summer 2020. This system is compatible with gDNA as a template and covers from within FR1 to the end of the J-gene.

To learn more about PER and sequencing depth, see our **guide on choosing read length/depth.**

CHAPTER FOUR

iRepertoire's sequencing services

COMPREHENSIVE AND SPECIALIZED
IMMUNE SEQUENCING

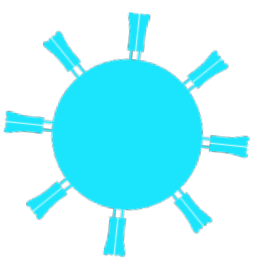
Immune repertoire sequencing: RepSeq and RepSeq+

iRepertoire's sequencing services **RepSeq** and **RepSeq+** complement our amplification technologies **arm-PCR** and **dam-PCR**, respectively. Because arm-PCR and dam-PCR have different strengths, the decision of whether to use RepSeq or RepSeq+ will depend on the nature of your specific project.

Arm-PCR uses internal and external nested primers for extremely inclusive repertoire sequencing; therefore, RepSeq is preferable for the discovery of low-frequency clonotypes, especially when cell counts are low. Arm-PCR can be performed in-house with standard laboratory equipment. For more information about arm-PCR products, please see our **products page.**

Dam-PCR achieves simultaneous amplification and sequencing of all 7 TCR and BCR chains. **RepSeq+** provides a more quantitative picture of the entire immune repertoire with unprecedented multiplexing abilities. Additionally, because unique molecular identifiers can be used in dam-PCR, RepSeq+ provides for more sensitive detection and correction of PCR and sequencing errors. When it is important to identify single nucleotide variability with high fidelity, as in B-cell receptor hypermutation analysis, RepSeq+ is a better choice. Dam-PCR must be completed as an iRepertoire service using RepSeq+, or in-house using the **iR-Flex automation system**.

For more information about arm-PCR and dam-PCR, see **Chapter Two: iRepertoire's amplification technologies**.



Batch sequencing

iRepertoire's amplification technologies and sequencing services are designed to be compatible with multiplexing so you can save costs by batch processing multiple samples into one sequencing run. If your order is not large enough to fill an entire flow cell, we can use other customer samples to fill the remaining space.

For example, let's say you wanted to study the TCR alpha and TCR beta repertoire of a single sample using our **RepSeq service**. This is equivalent to two barcoded libraries because two unique **reagent systems** are required for the amplification (one per chain). Sequencing would most likely be performed on the Illumina MiSeq with a V2-500 cycle kit, and because this kit can sequence up to 10 samples at a read depth of 1 million reads per sample, we would need to place eight other customer samples in your sequencing run to complete the flow cell.

Fewer samples (i.e., not filling an entire flow cell) could mean an extended time frame for sequencing to complete since the remaining space needs to be filled. Alternatively, if you have enough libraries to fill an entire sequencing lane, then this delay will not occur. If your libraries are multiplexed with other customer's libraries, we can provide the de-multiplexed data (not the raw sequencing data). This restriction does not apply if you fill an entire lane with only your samples.

For more information about multiplexing, see our page on **pooling samples**.

For more information about iRepertoires reagents, see the prior chapter about **iRepertoire's primer systems**.

CHAPTER FIVE

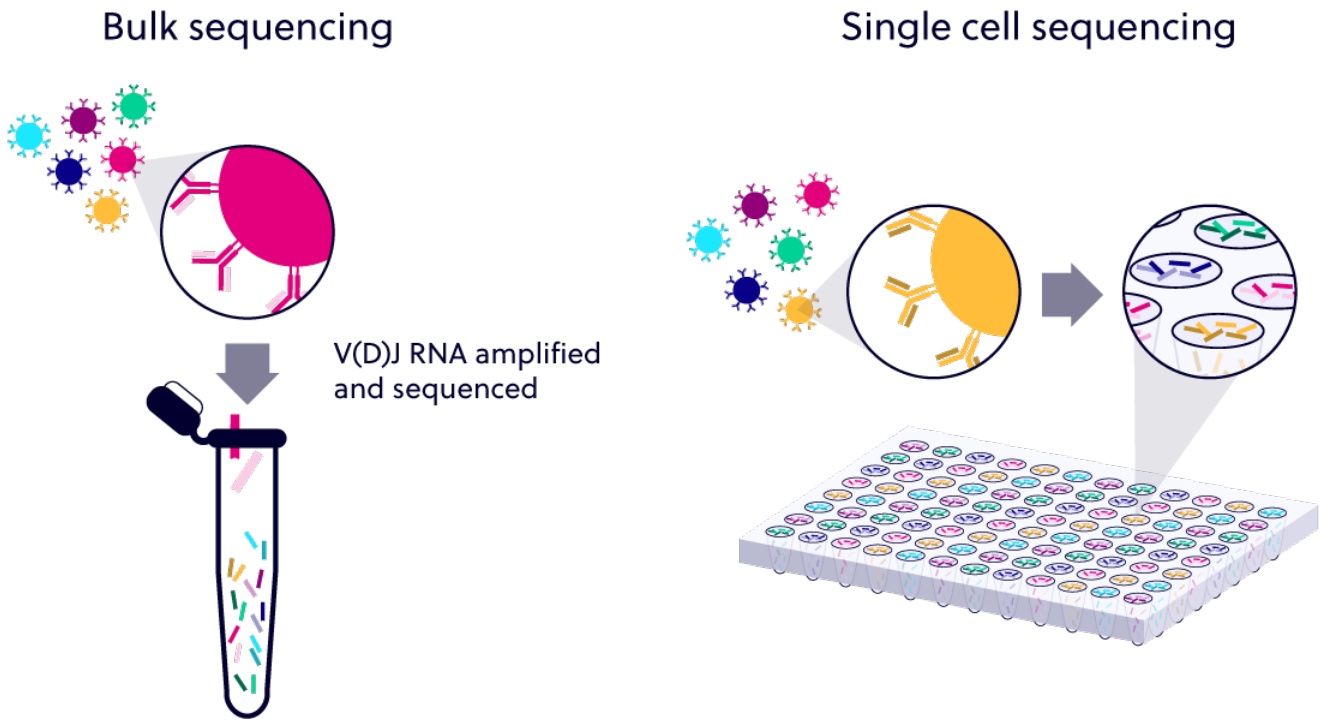
iPair: iRepertoire's single cell service

PRESERVE COGNATE PAIRING FOR DEEP AND ACCURATE INSIGHT

Why single cells?

Information related to the physical, cognate pairing of TCR-alpha and -beta chains or Ig-heavy and -light chains is lost once RNA extraction is performed on a bulk cell sample. iPair, our **single cell sequencing service**, was developed so that the repertoire of individual cells could be profiled through capture of chain pairing without the need to culture cells.

iPair was introduced on the market in November 2016. To our knowledge, iPair was the first consumer-based service specifically designed and intended for capturing the physically paired TCR alpha/beta or BCR IgH/KL variable region from human T and B cells. Since then, our services have expanded to cover mouse samples, as well as development of TCR gamma/delta pairing.



How it works

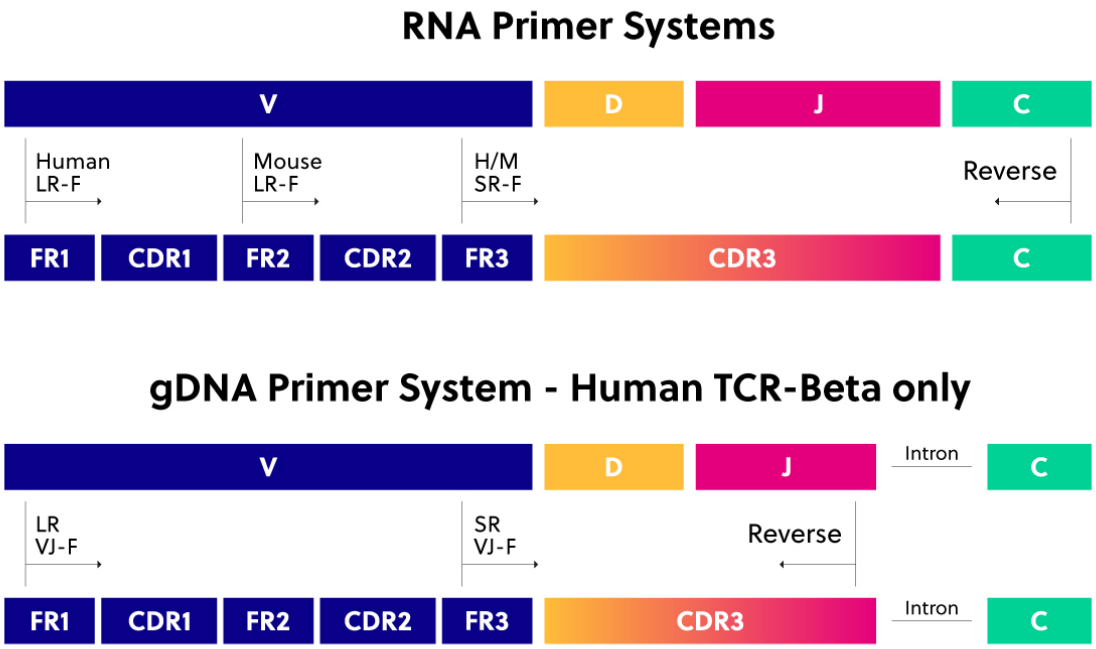
iPair relies on our arm-PCR technology. With iPair, we use a FACS-based approach to plate single cells into a 96-well iCapture plate. Our multiplex primer mixes were developed to balance the amplification of both loci simultaneously in the same PCR well in a low reaction volume. We then sequence the results using the Illumina MiSeq platform with 500+ cycle sequencing kits so that the variable region sequence can be called.

iPair generally achieves a **>90% amplification success rate**, and the pairing success rate is greater than 80% with our all-in-one targeted-seq mixes.

iPair primers

The iPair primers were designed using our long-read primer system, which covers from just within FR1 to the beginning of the C-gene for human samples (mouse long-read primers begin within FR2). Since we are only missing the first ~30-50 nucleotides of the V-gene, these sequences can be inferred from the reference sequences on the international ImMunoGeneTics information system (IMGT) quite easily. The placement of the primers produces libraries that are ideally sized for 500-cycle kits on Illumina platforms (250 paired-end reads). We have had many clients reconstitute functional TCRs from the data using the inferred sequences without issue as the majority of the TCR is present in the data including CDR1, CDR2, and CDR3.

To learn more about iRepertoire's primer systems, please see **Chapter Three on iRepertoire's primer system.**



All of iRepertoire's primer systems cover the highly variable CDR3 region

Single cell phenotyping with iPair+

Our **single cell phenotyping service**, iPair+, provides additional functional information about the cell(s) profiled using specialized primers. There are two different customizable options for iPair+, and which is best for your project largely depends on which phenotypic targets you'd like to call, how many targets you have, and your budget.

The gene-specific primer method involves co-amplification of phenotypic targets and the VDJ receptor primers in the same well. There are 30 gene-specific primers available for human samples; mouse samples must be processed with the oligo-dT approach.

The main advantage of using gene-specific primers is affordability. Because everything is processed in a single plate, there are no extra processing fees. However, highly expressed phenotypic targets could occupy a large amount of reads during sequencing, which could cause VDJ dropout. We recommend using the gene-specific approach if only a low number (~10-15) of phenotypic markers will be used.

Our second approach is an oligo-dT based method. The oligo-dT panel includes 100+ gene targets for either human or mouse samples. During this process, the cells are lysed and RT is performed with an oligo-dT primer mix. This cDNA is then split between two plates: one for VDJ amplification, and the other for phenotyping. Because the cell in each well is individually barcoded, phenotypic data can be overlaid with VDJ information during downstream data analysis. While this does increase the processing price, it allows for more user control of the sequencing depth of phenotypic targets.

Please **contact us** to view the gene-specific or oligo-dT panel lists.

Jump to the **iPair+ page** for specific information

Applications

New applications in immunotherapy and an increasing awareness of the importance of the tumor microenvironment have produced a growing demand for immune profiling at the single cell level. iPair can be used in CAR-T development, for calculating clonal frequency in cell subsets, or to track specific lymphocytes following treatment.

Contact us to learn more about how you can use iPair.

CHAPTER SIX

iRepertoire's data analysis services

COMPLIMETARY DATA ANALYSIS FOR INTUITIVE INSIGHT

Paired data analysis

All of our bulk repertoire services come with complimentary basic data analysis through our proprietary pipeline, which then outputs data to **iRweb**. iRweb is a web-based bioinformatic software that allows researchers to have all of the following services and analyses performed on their samples:

- The number of CDR3s captured in the library and the number of unique CDR3s within each sample
- The diversity of the immune repertoire in each sample, captured by a diversity index, a proprietary D50 value, and the Shannon entropy
- Normalized and unnormalized distributions of V/J-usage, V/J-trimming, CDR3 length, and N-addition. (Normalization accounts for differing coverage depth by treating each unique CDR3-VDJ combination as one regardless of read count.)
- VDJ-C mapping
- Alignments to the IMGT database
- V-J combination distributions as 2-D and 3D maps
- CDR3 hierarchical peptide frequency mapping including class-switching for BCRs
- CDR3 peptide with V-J frequency lists
- CDR3 algebra: Different samples are scaled, so that the frequency of CDR3s can be compared across samples with differing read depths.

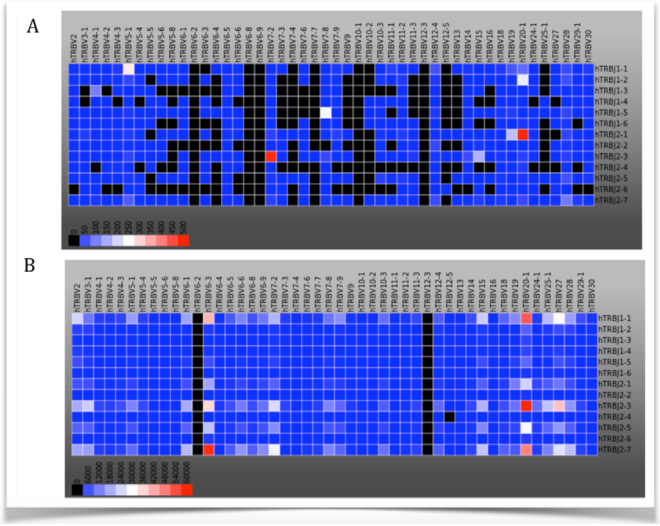
Access to iRweb and basic data analysis is also available to customers who have purchased our **reagent kits**, but certain rules and restrictions apply in order to avoid fees.

Learn more about iRweb and explore a demo via our **Bioinformatics page**.

Reports

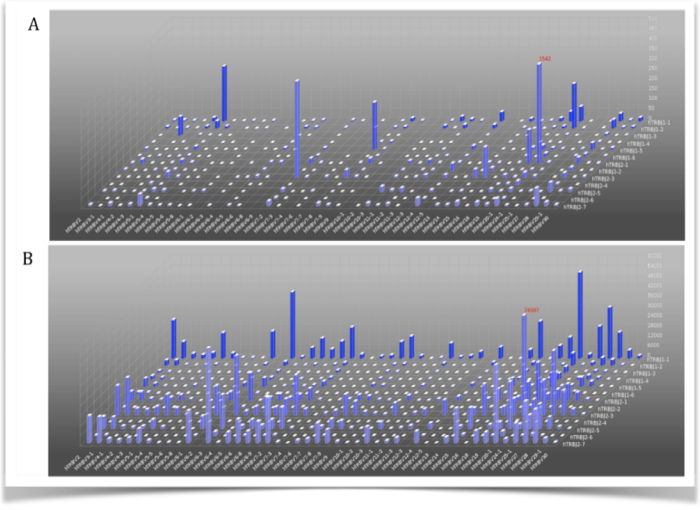
Heat Maps

The relative frequency of different alleles within a population is shown via heat maps. For instance, you might use a heat map to visualize what T-cell receptor gene variants appear most frequently in a healthy patient versus a patient with cancer. Quantitative differences in diversity (the number of different variants) are immediately obvious in 2D or 3D heat maps like the ones below:



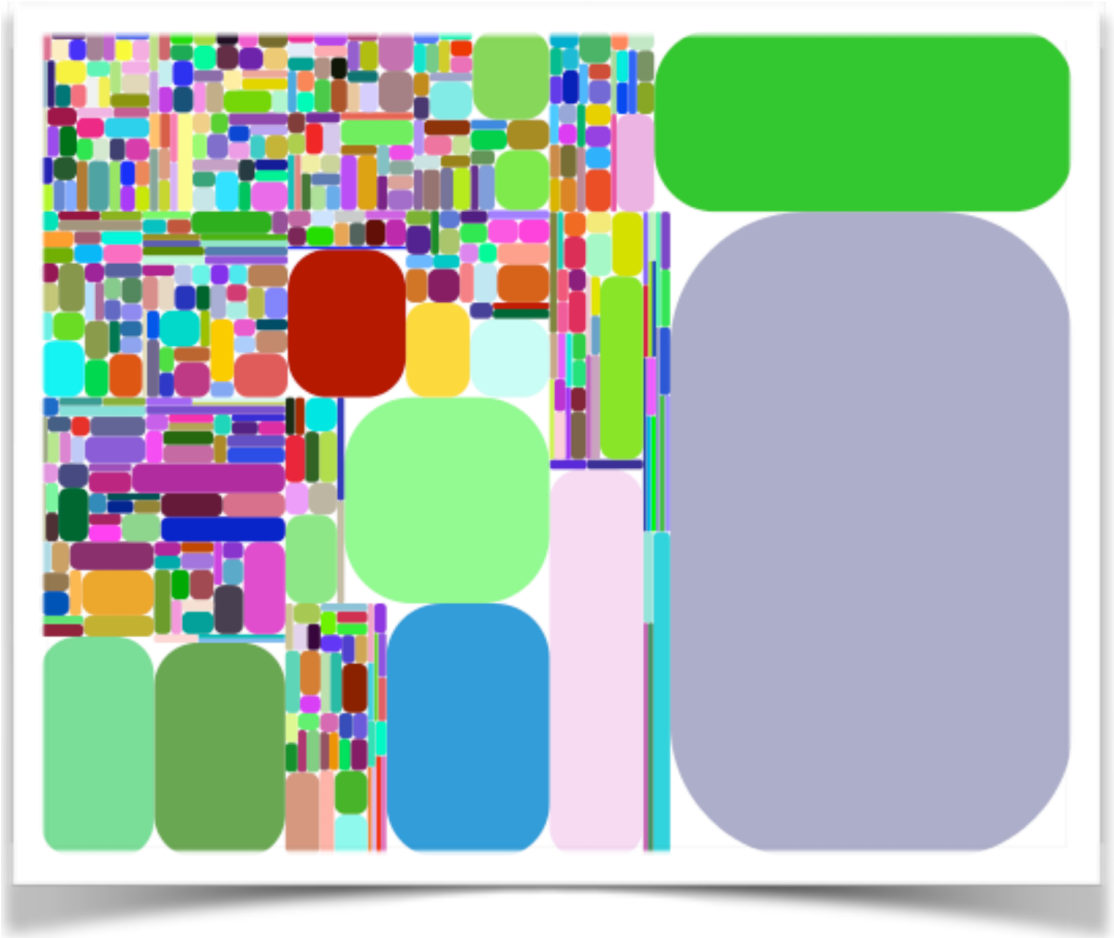
Above: 2D heat maps visualize the relative frequency of different alleles using color gradiants.

Below: 3D heat maps highlight differences in populations quickly and clearly, using differences in height.



Tree Maps

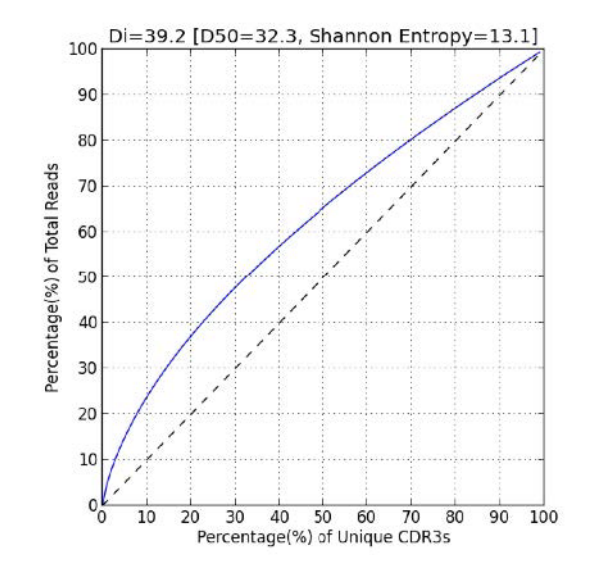
Diversity can also be shown using tree maps. Tree maps like the one below represent a variant population, in which each clonotype is represented by a different colored shape. The size of the shape reflects the frequency of the variant (i.e., a large shape represents a high frequency, or clonal expansion). Tree maps enable quick visual assessments of the relative repertoire diversity in different samples; smaller shapes and more varied colors equate to greater diversity.



Tree maps inherent to iRweb data analysis enable a quick visual assement of relative immune diversity in different samples. Smaller shapes and more varied color equal greater diversity.

Graphs

The diversity of the **immune repertoire** can also be captured numerically via a diversity index (Di), Shannon Entropy, and D50 value. The diversity index is a function of the frequency of each CDR3 and the total number of unique CDR3s. The Shannon entropy concerns only the top 10,000 most frequent CDR3s. The D50 is the percent of dominant and percent of unique clones that account for 50% of the total number of CDR3s in the sample. Very diverse libraries will have a D50 value close to 50. The diversity is represented by graphs like the one below in which the black dashed line represents "perfect" diversity.



Shannon entropy graph

iPair Analyzer

The **iPair Analyzer** is the data analysis platform developed for **single cell repertoire sequencing** data performed via our iPair service. The iPair Analyzer aids in visualization and comparison of data for single cells as well as bulk repertoires (if bulk sequencing is performed in tandem).

The iPair Analyzer's exclusive graphic user interface represents samples in a 96 well plate in an interactive panel. Features such as read depth, chain type, or the presence of a particular unique CDR3 can be highlighted via different colors and symbols on the panel. When a cell in the panel is selected, single cell BCR or TCR chain results are displayed. A worksheet feature lets you select chains of interest and further explore and compare additional features such as the nucleotide sequence of the read, CDR1, CDR2, CDR3, V, D, and J usage.

iPair Analyzer view example

NEXT STEPS

LEARN MORE

Continue your learning journey to extend knowledge of NGS immune repertoire sequencing, the adaptive immune system, and innovative new directions in immune research.

Find more education and research resoureces and tools in the online Learning Center at iRepertoire.com.

START NOW

iRepertoire, Inc.