



TechNote

Understanding the immunoSEQ Assay: **From Inquiry to Insights**

For Research Use Only. Not for use in diagnostic procedures.



INTRODUCTION

The Adaptive Biotechnologies Corp. proprietary platform for immunosequencing, the immunoSEQ™ Assay, combines novel multiplex PCR¹ with highly optimized primer sets and deep sequencing techniques that exclusively target T-cell receptor (TCR) and B-cell receptor (BCR) genes. The immunoSEQ platform enables researchers to analyze the adaptive immune system with exceptional depth and specificity.

To fully utilize the potential of the immunoSEQ Assay, it is important to understand the breadth and depth of information the immunoSEQ Assay yields. This overview provides background on the immune repertoire and highlights the various types of results generated by the immunoSEQ Assay.

The immunoSEQ platform enables researchers to analyze the adaptive immune system with exceptional depth and specificity

THE IMMUNE REPERTOIRE

The adaptive immune system is composed of T and B lymphocytes. T lymphocytes express TCRs on their surface to recognize antigenic peptides presented by host cells. B lymphocytes produce antibodies (BCRs) that bind to antigens outside of the cell. To generate a receptor repertoire with sufficient diversity to recognize the vast number of potential antigens, T and B lymphocytes assemble functional receptors. The T and B lymphocytes combine noncontiguous variable (V), diversity (D), and joining (J) region gene segments—which collectively encode the complementarity determining region 3 (CDR3)—through a process that involves rearrangement of germline DNA. Diversity of the CDR3 region is augmented by the deletion and template-independent insertion of nucleotides at the V-D and D-J junctions, and further by somatic hypermutation in the BCR. The total number of CDR3 possibilities in a particular individual is termed the ‘immune repertoire’.

The total number of CDR3 possibilities in a particular individual is termed the ‘immune repertoire’

The inherent diversity of repertoire formation makes it highly improbable to independently create the same CDR3 nucleotide sequence twice, effectively making each CDR3 sequence a unique nucleotide tag for any specific T or B cell and its clonal descendants. The rearrangement of the CDR3 region occurs after cell lineage commitment, and only in lymphocyte populations; this unique nucleotide tag is the marker the immunoSEQ Assay uses to track T and B clones over time and to assess clonality and diversity of lymphocyte populations.

Blood samples can be sequenced to offer a snapshot of the immune repertoire in the periphery. Samples can also come from individual tissues or sorted T- or B-cell sub-populations to give a more focused view of the repertoire. Pairing different sample types, such as peripheral blood and tissue, is a good approach to understanding the repertoire in a particular disease setting.

Each CDR3 sequence is a unique nucleotide tag for any specific T or B cell and its clonal descendants

Most receptor sequences will be found in only one or a few cells in an individual. During the formation of immunological memory, however, cells undergo clonal expansion. Consequently, after expansion and conversion to memory cells, receptor sequences against past pathogens can be present in thousands of cells. The large potential diversity of receptor sequences means that in

most cases, nucleotide-identical sequences are not shared between cells except through clonal expansion.

By quantitating the exact number of input cells that contribute to an observed sequence, the immunoSEQ Assay allows researchers to analyze features of the highly expanded clones and also of infrequent cells. A sampling of the immune repertoire cannot capture everything; sampling all cells in 1 milliliter of blood is only a sample of approximately 0.02% of the repertoire in the periphery. This is itself only a small subset of all immune cells, but expanded clones will be identified reproducibly.

THE POWER OF IMMUNE REPERTOIRE ANALYSIS ENABLED BY THE IMMUNOSEQ ASSAY

Immune repertoire samples can be processed with the immunoSEQ Assay either by purchasing the immunoSEQ Human T-Cell Receptor Beta Kit (for research use only) or by shipping samples directly to Adaptive Biotechnologies. After processing, you will receive a wealth of data from your samples as well as access to the online immunoSEQ Analyzer. The immunoSEQ Analyzer is a robust and user-friendly toolset to help you visualize the broader contours and dimensions of your data, to better understand its implications, and get to new discoveries faster. Key assessments enabled by the immunoSEQ Analyzer include clone tracking, repertoire characterization, and public clone identification.

Key assessments include clone tracking, repertoire characterization, and public clone identification

Data generated by the immunoSEQ Assay generally fall into one of three categories:

1. Sequence-level results describe characteristics of individual TCR or BCR sequences
2. Sample-level results summarize a sample's entire repertoire
3. Project-level results combine information from the sequence and sample levels with other metadata from the study design such as sample identification, case or control tags, and time-course information, to give cross-sample and cross-subject insights

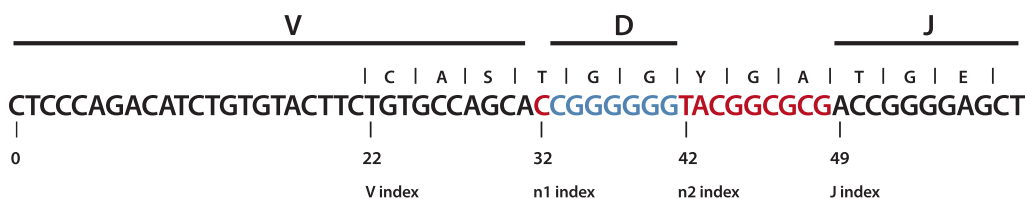
ANALYSES FOCUSED ON SINGLE SEQUENCES

Sequence-level results include:

- Relative frequency of top clone(s)
- CDR3 length and amino acid sequence
- V-, D-, and J-gene identity

Sequence-level analysis describes characteristics of a single clone, or unique nucleotide sequence. Sequencing reads belonging to the same clone are identified by having perfectly matched nucleotide sequences in the CDR3 region. For a clone to be included in the results, its unique nucleotide sequence must be identified by the sequencer at least twice. The data output includes counts, indicating how many identical reads were measured. The total count of reads for a given clone represents the relative size of that clonal population in the repertoire. Sequencing depth (the total number of reads per sample) can vary; therefore, all reads are normalized and the frequency (the percentage of reads counted by total number of reads) of each clone is reported to allow cross-sample comparisons.

Adaptive Biotechnologies' proprietary immunoSEQ Analyzer allows the user to view detailed information about each clone, starting with the translated amino acid sequence. An amino acid sequence is identified as productive (fully functional); out of frame (disrupted translation frame due to a CDR3 length that is not a multiple of three); or containing a stop (premature stop codon in sequence, rendering the sequence nonfunctional). The length of the reported sequence depends on the assay in use, but averages 100 bp, which is enough for identification of the V, D, and J genes and for the overall CDR3 length (Figure 1). Information on the CDR3 length can be used to compare the immunoSEQ Assay results with spectratyping results.



V, variable; D, diversity; J, joining.

Figure 1. Overview of the CDR3 region, including the V-, D-, and J-gene segments and random nucleotide insertions (in red) identified by the immunoSEQ Assay.

Our sophisticated primer sets as a target the highly variable CDR3 region of the heavy chain of immunoglobulins that contains a fraction of the V region, the complete D region, and the J region.² The identification of specific V, D, and J genes is carried out according to the definitions established by the International ImMunoGeneTics (IMGT®) collaboration.³ Our IGH assay for B cells also captures incomplete DJ rearrangements in parallel that are filtered out by default in the immunoSEQ Analyzer, but are available for each sample when desired.

Due to the higher resolution of high-throughput DNA sequencing, the immunoSEQ Assay provides more information than can be identified via traditional flow cytometry. To facilitate comparison, the corresponding flow cytometry nomenclature is provided for all V, D, and J classifications where available. For some V, D, and J classifications, there is no corresponding flow cytometry antibody. The immunoSEQ Analyzer includes a tool for making a V gene vs. sample percentage chart. This chart differentiates the V genes identifiable by both flow cytometry and sequencing, from the V genes identifiable only with the immunoSEQ Assay (Figure 2).

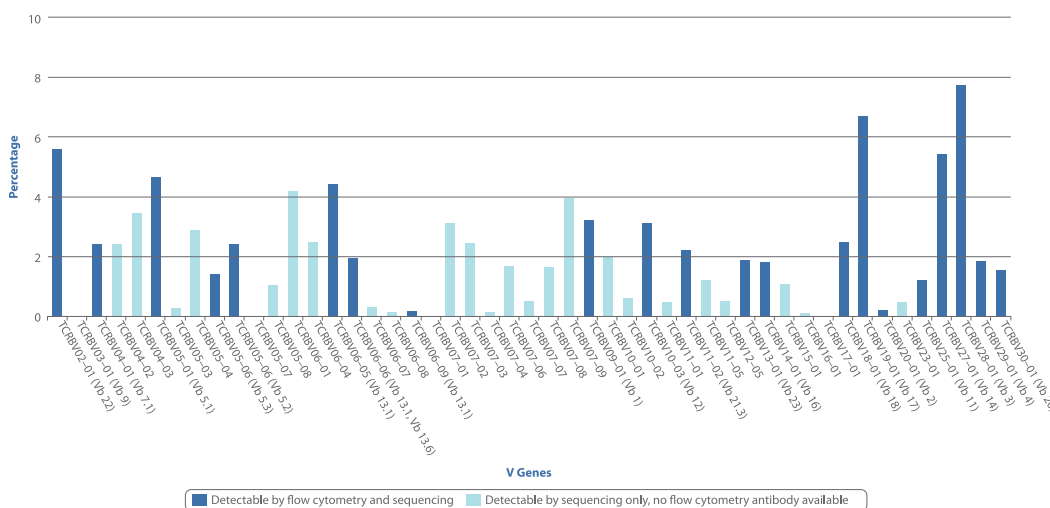


Figure 2. Screenshot of the V Mark tool. It differentiates clones that can be identified by both spectratyping and sequencing (dark blue) and clones that can be identified only by sequencing with the immunoSEQ Assay (light blue).

Sometimes the specific V, D, or J gene in a clone's sequence cannot be identified unambiguously. In such a case, the software lists the V, D, or J family or gene as unresolved and indicates a narrowed set of potential V, D, or J classifications elsewhere in the results.

ANALYSES FOCUSED ON SINGLE SAMPLES

The immunoSEQ Analyzer facilitates the user examining global characteristics of a sample's repertoire.

Sample-level results include:

- Clonality—a statistic for how much of the repertoire is made up of expanded clones
- Diversity—an estimate of how many clones exist in the entire repertoire
- Total unique sequences/total cell count
- Other aggregated statistics such as overall V/D/J usage or overall CDR3 length—the immunoSEQ Analyzer allows aggregation over all reads (using 'total count'), or a flattened picture of the repertoire before clonal expansion (using 'unique count') where each unique sequence is treated equally

Adaptive Biotechnologies' sequencing results describe the repertoire as the entire set of measured clones. The term clonality in the results describes the degree to which one or a few clones dominate the repertoire. Adaptive Biotechnologies' calculated measure of clonality varies from 0 to 1, with 0 being a flat distribution (each clone appears only once) and 1 being an entirely oligoclonal sample (only one clone was found and all reads come from that one clone). It is the inverse of the normalized version of Shannon's entropy⁴ and is a measurement of the shape of the frequency distribution of all productive clones in the sample.

While clonality is reported as a single number for a sample, an intuitive view of clonality can be achieved by looking at the relative proportions of the top clones in a sample, as in Figure 3.

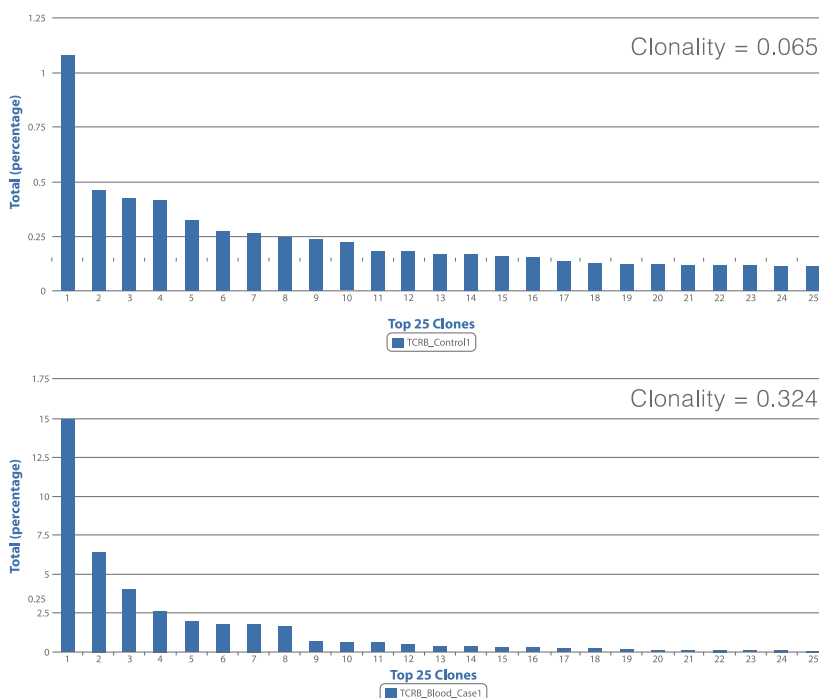


Figure 3. Screenshot of the Top Clones tool for two samples with differing clonality.

Only one clone in the top sample is above 1% frequency. In comparison, in the higher clonality sample on the bottom, the largest clone is around 20% frequency. Several others are above 1%.

Due to the highly standardized methods and controls included, the immunoSEQ Assay also generates data on the percentage of B and T cells in a given sample. When the total amount of input genomic DNA (gDNA) is also known, users can calculate the number of B and T cells in the original sample. This kind of cell-counting data is particularly useful when analyzing the immune repertoire in tissue samples.

The immunoSEQ Analyzer facilitates examining global characteristics of a repertoire like diversity and clonality

ANALYSES FOCUSED ON MULTIPLE SAMPLES

When samples have been collected from multiple tissues, time points, or individuals, the ImmunoSEQ Analyzer facilitates project-level analyses.

Project-level results include:

- Subject-specific analysis, such as clone tracking, which can be an important research tool for understanding the temporal diversity of the adaptive immune system. This answers questions such as, “In this subject, which clones expanded from time point A to time point B?”
- Cross-subject analysis, such as identifying common sequences. This answers questions such as, “Which sequence/sample-level features are different in my case vs. control groups?” or “Are there any public clones seen in multiple individuals?”

These Analyzer capabilities make it possible to see the effects of new and unknown vaccinations, the long-term effects of drugs and therapies, or the decline of immune cells during chemotherapy in the research setting.⁵ The immunoSEQ Assay also enables scientists to identify antigen-specific receptor sequences and to determine the specificity of cells.

One clear example of a project-level analysis is the Clone Tracker tool (Figure 4). Any clone of interest in a specific sample can be viewed in other samples to see how its relative frequency has changed. The Clone Tracker can identify expanded clones that were part of a subject’s immune response to a foreign antigen.

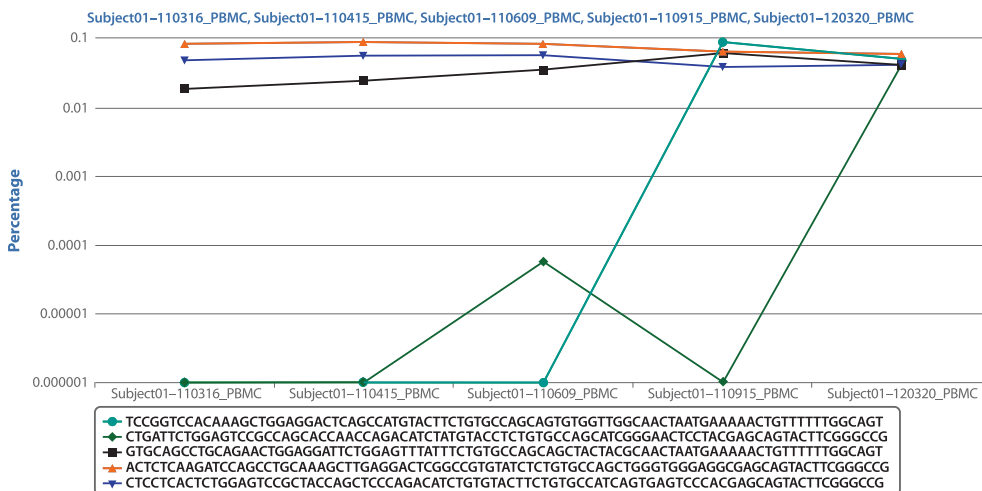


Figure 4. Clone Tracker example for a project with five time points.

Large expanded clones (in differing shades of blue) in the first time point are seen consistently at other time points. Other infrequent clones from the first time point (in gray) or not in the first time point at all (in pink) are shown at much lower frequencies. Two of the tracked clones expand between months 3–6 and months 6–12, respectively.

Another project-level tool is the Common Sequences tool that enables rapid identification of shared clones between samples. By comparing receptor sequences from different individuals, the immunoSEQ Analyzer can help to identify public clones.

By comparing receptor sequences from different individuals, the immunoSEQ Analyzer can help to identify public clones

SUMMARY

The immunoSEQ Analyzer allows researchers to examine three detail levels (sequence, sample, and project) to identify trends in their data. A well-designed experiment requires consideration of all of these levels in order to maximize the utility of the resulting information. By enabling assessments like clone tracking, public clone identification, and repertoire characterization, the immunoSEQ platform empowers researchers to better understand the body's response to disease at the cellular level.

To speak with one of Adaptive Biotechnologies' scientific advisors about your experiment or to place an order, call 1-855-466-8667 or visit www.adaptivebiotech.com.

References

1. Carlson CS, Emerson RO, Sherwood AM, et al. Using synthetic templates to design an unbiased multiplex PCR assay. *Nat Commun* 2013;4:2680.
2. Robins HS, Campregher PV, Srivastava SK, et al. Comprehensive assessment of T-cell receptor beta-chain diversity in alpha beta T cells. *Blood* 2009;114:4099–107.
3. Yousfi Monod M, Giudicelli V, Chaume D, Lefranc MP. IMGT/JunctionAnalysis: the first tool for the analysis of the immunoglobulin and T cell receptor complex V-J and V-D-J JUNCTIONS. *Bioinformatics* 2004;20(Suppl. 1):i379–85.
4. Emerson RO, Sherwood AM, Rieder MJ, et al. High-throughput sequencing of T-cell receptors reveals a homogeneous repertoire of tumour-infiltrating lymphocytes in ovarian cancer. *J Pathol* 2013;231:433–40.
5. Wu D, Sherwood A, Fromm JR, et al. High-throughput sequencing detects minimal residual disease in acute T lymphoblastic leukemia. *Sci Transl Med* 2012;4:134ra63.

For research use only. Not for use in diagnostic procedures.

Copyright © 2014 Adaptive Biotechnologies Corp. All rights reserved. 10/2014

Adaptive Biotechnologies and immunoSEQ are trademarks of Adaptive Biotechnologies Corporation.

All other trademarks are properties of their respective owners.



1551 Eastlake Avenue East, Suite 200 • Seattle, WA 98102

P: 855.466.8667 **W:** adaptivebiotech.com

MRK-10013 AA