APPLICATION NOTE

Full-length V(D)J IR-Seq on the NextSeq™ 1000 and NextSeq 2000 Systems

- More data with higher quality reads allows discovery of more clonotypes
- More samples per run for flexible experimental designs to meet research budgets

illumina®

Immune repertoire biology

The adaptive immune system, made up of B lymphocytes and T lymphocytes, provides highly specific protection from pathogens through an evolutionarily refined process that produces a massively diverse repertoire of antigen receptors.[1] When an antigen receptor recognizes a pathogen, a series of events takes place to amplify that receptor and recruit additional immune cells to eliminate the infection. This process for generating receptor diversity is known as V(D)J recombination. In V(D)J recombination, variable (V), diversity (D), and joining (J) segments are stochastically recombined into functional B-cell receptors (BCRs) and T-cell receptors (TCRs). Genes for BCRs and TCRs occur at different sites in the genome as arrays of variable V, D, and J segments, as well as constant regions that make up the rest of the receptor structure (Figure 1 and Figure 2). By randomly recombining V, D, and J segments from a small number of genes into mRNA transcripts, it is estimated that it is possible to generate 10^{18} unique BCR sequences[2,3] and 10^{12} unique TCR sequences.[4,5]

The quantifiable immune repertoire of BCR and TCR sequences is dynamic in response to infections, autoimmune disorders (eg, lupus, rheumatoid arthritis, multiple sclerosis), and cancers. During an immune response, the repertoire of circulating antigen receptors shifts from a diverse pool to one that is dominated by on or a few expanded clones (Figure 3). Therefore, understanding the makeup of the repertoire is highly informative for what is happening in the disease process. The incredible diversity of the immune repertoire means that analysis requires high sequencing depth. In addition, accurate assembly of the full-length recombined V(D)J genes plus partial constant region containing isotype-determining information that ranges from 400 bp to 600 bp[6] requires a platform capable of 2 × 300 paired-end reads to form a single, gapless contig for the variable region sequence reconstruction.

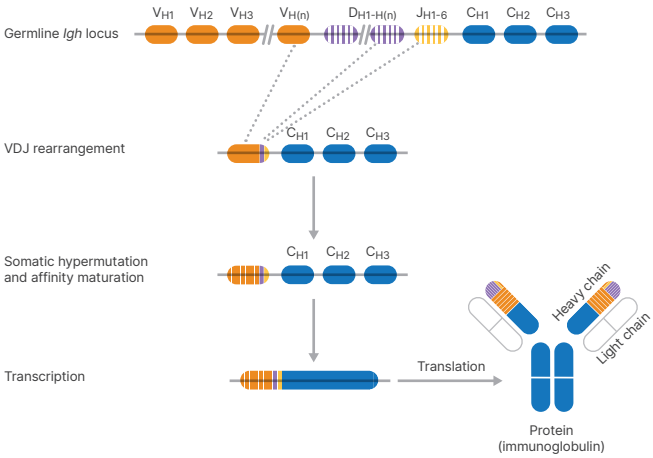

Figure 1: BCR generation through the V(D)J recombination process—Genes for BCR receptors occur at multiple sites in the genome. The genes contain a large number of V domains and a smaller number of D and J domains, as well as constant C domains. Stochastic recombination of some of the V(D)J domains results in a highly diverse population of immunoglobulin receptors.

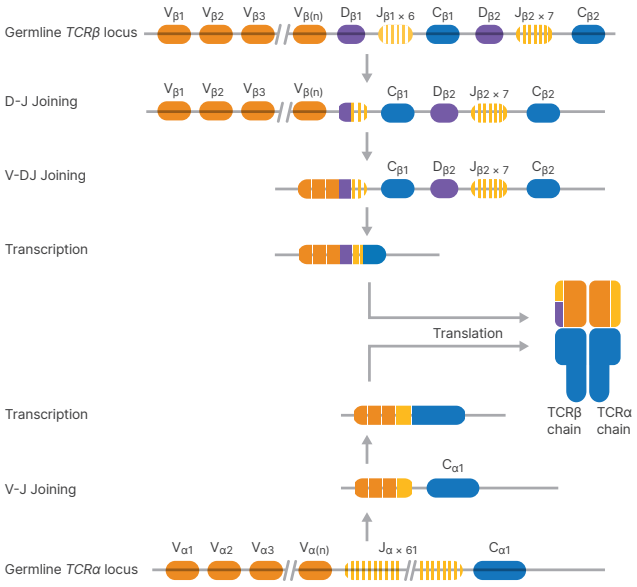

Figure 2: TCR generation through the V(D)J recombination process—TCRs are heterodimers transcribed from two different genomic loci. The TCRβ gene contains several V, D, J, and constant domains while the TCRα contains V, J, and constant domains. Stochastic recombination of V, D, and J domains results in a highly diverse population of TCRs.

The availability of 600-cycle kits makes the NextSeq 1000 and NextSeq 2000 Systems an excellent choice for labs looking for a powerful benchtop sequencer for immune repertoire sequencing (IR-Seq) and other NGS applications. The NextSeq 1000/2000 P1 XLEAP-SBS™ Reagent Kit (600 cycles) (Illumina, Catalog no. 20100981) generates up to 60 Gb of high-quality data with 100M reads and the NextSeq 1000/2000 P2 XLEAP-SBS Reagent Kit (600 cycles) (Illumina, Catalog no. 20100984) generates up to 240 Gb of high-quality data with 400M reads. In comparison, the

MiSeq™ Reagent Kit v3 (600-cycle) (Illumina, Catalog number MS-102-3003) generates up to 15 Gb of high-quality data with 25M reads.

IR-Seq library preparation

The choice of an immune repertoire library solution depends on various factors such as input (ie, DNA, bulk RNA, or single-cell RNA), methodology (ie, rapid amplification of cDNA ends or multiplex PCR), regions sequenced (ie, complementarity determining region (CDR3) or full-length), and the chain type (ie, TCR or BCR).[7] Multiple library prep solutions and protocols for immune repertoire studies are available and several have been tested on the NextSeq 2000 System (Table 1). Many of these options require, or benefit from, 2 × 300 paired-end sequencing. NextSeq 1000/2000 XLEAP-SBS Reagent Kits (600 cycles) represent an advanced alternative to the MiSeq Reagent Kit v3 (600-cycle) for researchers looking to scale their experiments.

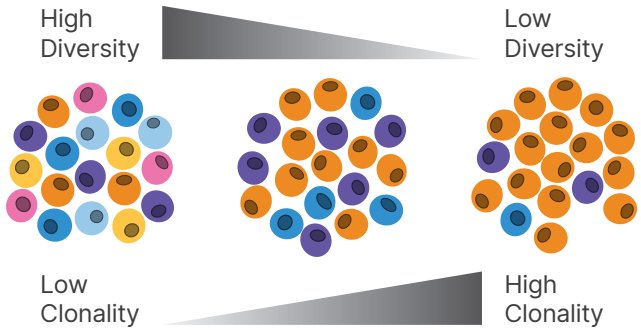

Figure 3: Adaptive immune response—The adaptive immune system relies on a highly diverse population of cells expressing unique BCRs and TCRs. In the absence of an antigen response, the repertoire shows high diversity and low clonality of receptor types. In the presence of a recognized antigen, cells with activated receptors are expanded, resulting in low diversity in the repertoire and high clonality of the active cells.

Table 1: Tested third-party library prep kits for IR-Seq applications[a,b]

Resolution	Method	Provider	Kit name (Catalog no.)	Targeted receptors	Sequencing read length recommendation	Bioinformatics secondary analysis
Bulk RNA-Seq	CDR3/full-length	New England Biolabs	NEBNext Immune Sequencing Kit (Human) (E6320S, E6320L)	BCR, TCR or BCR + TCR	2 × 300 bp	Open Source pRESTO Tools
Bulk RNA-Seq	CDR3/full-length	New England Biolabs	NEBNext Immune Sequencing Kit (Mouse) (E6330S, E6330L)	BCR, TCR or BCR + TCR	2 × 300 bp	Open Source pRESTO Tools
Bulk RNA-Seq	CDR3/full-length	QIAGEN	QIAseq Immune Repertoire RNA Library Kit (333705)	TCR	2 × 300 bp	QIAGEN GeneGlobe pipeline
Bulk RNA-Seq	CDR3/full-length	Takara	SMART-Seq Human TCR (with UMIs) (634780, 634781, 634779)	TCR	2 × 300 bp	Cogent NGS Immune Profiler Software
Bulk RNA-Seq	CDR3/full-length	Takara	SMART-Seq Human BCR (with UMIs) (634777, 634778, 634776)	BCR	2 × 300 bp	Cogent NGS Immune Profiler Software
Single cell	CDR3/full-length	BD Bio	BD Rhapsody TCR/BCR Multiomic Assay Kit (665828, 665829)	BCR, TCR or BCR + TCR	85 × 215 bp[c] or 2 × 300 bp	BD Rhapsody Analysis Pipeline

a. Third-party kits have been tested using NextSeq 1000/2000 Reagents (600 cycles) with standard SBS chemistry. XLEAP-SBS chemistry is now available for P1, P2, P3, and P4 reagent kits.
b. Represented libraries use unique molecular index (UMI) technology, providing error correction, deduplication, and high confidence reads coming out of the pipeline.
c. Run configurations for single cell are compatible with a smaller 300-cycles kit cartridge.

IR-Seq performance

IR-Seq libraries typically have large fragment sizes (600–800 bp) and a limited capability for multiplexing due to read depth requirements. The NextSeq 1000/2000 P1 XLEAP-SBS Reagent Kit (600 cycles) can generate 100M reads, or approximately 4× more data than the MiSeq System using the MiSeq Reagent Kit v3 (600-cycle). The NextSeq 1000/2000 P2 XLEAP-SBS Reagent Kit (600 cycles) generates an impressive 400M reads, or 16× more data versus the MiSeq System using the MiSeq Reagent Kit v3 (600-cycle). This substantial data capacity on the NextSeq 1000 and NextSeq 2000 Systems adds flexibility for increased read depth over the MiSeq System and for multiplexing of IR-Seq libraries (Table 2).

The NextSeq 1000/2000 P1 XLEAP-SBS Reagent Kit (600 cycles) and NextSeq 1000/2000 P2 XLEAP-SBS Reagent Kit (600 cycles) on the NextSeq 2000 System also produce a higher percentage of reads with quality scores ≥ Q30 compared to the MiSeq Reagent Kit v3 (600-cycle) on the MiSeq System. The improved Q30 scores across the length of the reads allow for better secondary analysis results and better clonotype identification (Table 2).

Finally, sequencing on the NextSeq 1000/2000 P1 XLEAP-SBS Reagent Kit (600 cycles) is significantly faster than the MiSeq System. The time savings could allow labs to complete three sequencing runs per week with the NextSeq 1000/2000 P1 XLEAP-SBS Reagent Kit (600 cycles), with more data per run, compared to two runs per week using the MiSeq Reagent Kit v3 (600-cycle) (Table 2).

Sequencing data generated with the NextSeq 1000 and NextSeq 2000 Systems can be output in standardized file formats that are compatible with an extensive ecosystem of commercial and open-source bioinformatic analysis tools. Each library prep provider offers their own custom analysis. Also, third-party dedicated tools are available for immune repertoire analysis.

Summary

The state of the adaptive immune system is directly tied to the health of an individual. Therefore, understanding the sequence composition of the immune repertoire provides researchers with a wealth of information about the health of a person, including how they are responding to an infection, the health of their immune system, and even the progress of certain cancers.

NextSeq 1000 and NextSeq 2000 Systems are compatible with a wide range of library preparation kits and analysis software from Illumina as well as third-party suppliers. With the available 600-cycle kits, these systems are excellent for complex IR-Seq applications and provide the reads necessary for a detailed view of the immune repertoire and the capacity for multiplexing of samples. Different flow cell configurations allow researchers to adjust read depth and sample numbers to match their experimental needs. The NextSeq 1000 and NextSeq 2000 Systems offer extensive cross-application flexibility, enabling researchers to transition easily between sequencing projects.

Table 2: NextSeq 1000, NextSeq 2000, and MiSeq Systems 600-cycle kit performance

	MiSeq Reagent Kit v3 (600-cycle)	NextSeq 1000/2000 XLEAP-SBS P1 Reagent Kit (600 cycles)	NextSeq 1000/2000 XLEAP-SBS P2 Reagent Kit (600 cycles)
Quality scores (percentage of bases ≥ Q30)	≥ 70%	≥ 85%	≥ 85%
Maximum total read pairs per run	25M	100M	400M
Maximum sample read pairs per run[a]	22.5M	90M	360M
Sample multiplexing to reach 5M reads[a]	4	18	72
Sample multiplexing to reach 25M reads[a]	–	3	14
Run times	55 hr	34 hr	42 hr

a. Maximum sample reads when using 10% (v/v) PhiX library addition for sequence diversity.

Learn more

NextSeq 1000 and NextSeq 2000 Systems

Immune-Related Genetic Variation

QIAGEN QIAseq Immune Repertoire RNA Library Kit data on the NextSeq 2000 600-cycles kit

New England Biolabs NEBNext Immune Sequencing Kit data now available on the new NextSeq 2000 600-cycles kit

Takara SMARTer Human BCR data now available on new 600-cycles kit on the NextSeq 2000 System

References

1. InformedHealth.org. Cologne, Germany: Institute for Quality and Efficiency in Health Care (IQWiG); 2006-. The innate and adaptive immune systems. https://www.ncbi.nlm.nih.gov/books/NBK279396/. Updated July 30, 2020. Accessed May 15, 2023.
2. Elhanati Y, Sethna Z, Marcou Q, Callan CG Jr, Mora T, Walczak AM. Inferring processes underlying B-cell repertoire diversity. *Philos Trans R Soc Lond B Biol Sci*. 2015;370(1676):20140243. doi:10.1098/rstb.2014.0243
3. Hoehn KB, Fowler A, Lunter G, Pybus OG. The Diversity and Molecular Evolution of B-Cell Receptors during Infection. *Mol Biol Evol*. 2016;33(5):1147-1157. doi:10.1093/molbev/msw015
4. Elhanati Y, Murugan A, Callan CG Jr, Mora T, Walczak AM. Quantifying selection in immune receptor repertoires. *Proc Natl Acad Sci U S A*. 2014;111(27):9875-9880. doi:10.1073/pnas.1409572111
5. Laydon DJ, Bangham CR, Asquith B. Estimating T-cell repertoire diversity: limitations of classical estimators and a new approach. *Philos Trans R Soc Lond B Biol Sci*. 2015;370(1675):20140291. doi:10.1098/rstb.2014.0291
6. Kim D, Park D. Deep sequencing of B cell receptor repertoire. *BMB Rep*. 2019;52(9):540-547. doi:10.5483/BMBRep.2019.52.9.192
7. Frank ML, Lu K, Erdogan C, et al. T-Cell Receptor Repertoire Sequencing in the Era of Cancer Immunotherapy. *Clin Cancer Res*. 2023;29(6):994-1008. doi:10.1158/1078-0432.CCR-22-2469

illumina®

1.800.809.4566 toll-free (US) | +1.858.202.4566 tel
techsupport@illumina.com | www.illumina.com

M-GL-01149 v2.0