A New Scalable Tool to Enable Full-Length V(D)J Immune Repertoire Sequencing (IR-Seq) For Cancer Research

Robin Bombardi[1], Qiyuan Han[2], Michael Morikado[2], Tinya Fleming[2], Sujash Chatterjee[2], Sophie Wehrkamp-Richter[2], Diana Lok[1], Jinghua Shi[2], Dalia Daujotyte[2], and Gary Schroth[1]

[1]Illumina, Inc, Foster City, CA; [2]Illumina, Inc, San Diego, CA

INTRODUCTION

The role of the adaptive immune system, made up of B lymphocytes and T lymphocytes, has gained notable attention in the last few years. Immuno-therapies such as immune checkpoint inhibitors (Figure 1) and CAR-Ts have shown promising results across multiple cancer types[1]. Moreover, immune repertoire analysis (sequencing of B and T antigen receptors) has been shown to be a potential useful tool for biomarker identification and prognostic value for the disease.

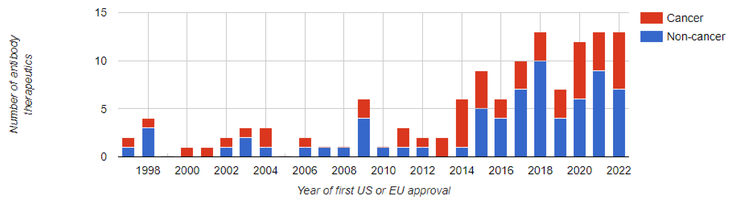

Figure 1: Number of antibody therapeutics granted a first approval in either the US or EU each year, 1997-2022

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 2a). Through V(D)J recombination, it is estimated that it is possible to generate 10^{18} unique BCR sequences[2] and 10^{12} unique TCR sequences[3]. However, during an immune response, the repertoire of circulating antigen receptors shifts from a diverse pool to one that is dominated by one, or a small number of expanded clones (Figure 2b).

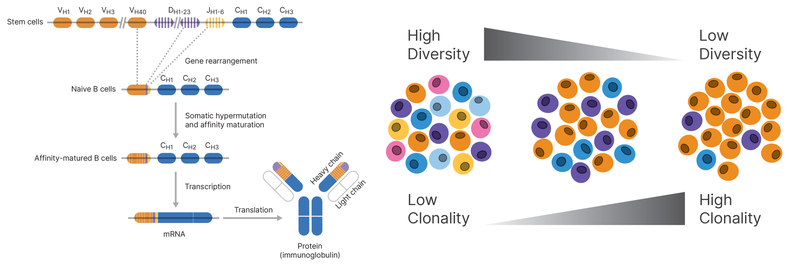

Figure 2: a) Mechanism of V(D)J recombination; b) dynamics of immune repertoire showing how in the absence of an antigen response the repertoire shows high diversity and low clonality of receptor types. In the presence of a recognized antigen, cells with activated receptors are expanded resulting in low diversity in the repertoire and high clonality of the active cells.

CHALLENGES & GOALS

- The incredible diversity of the immune repertoire means that analysis requires very high sequencing depth
- In addition, unambiguous, full-length, recombined, V(D)J-receptor reconstruction often requires or greatly benefits from 2X300 reads
- Here we demonstrate that the NextSeq™ 1000 and NextSeq 2000 600-cycle kits are compatible with a wide range of library preparation kits and analysis software from third-party suppliers

NEB blogpost | Qiagen blogpost | Takara blogpost

SEQUENCING & LIBRARY PREP SOLUTIONS

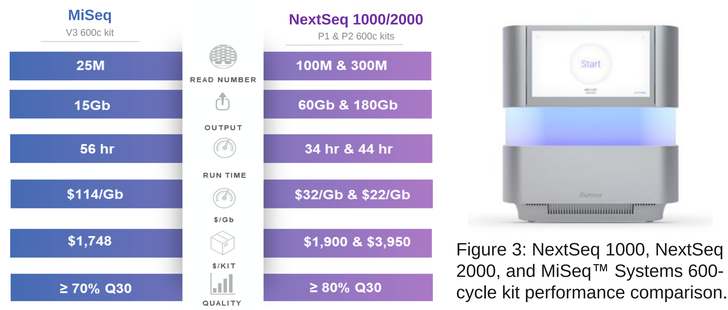

Figure 3: NextSeq 1000, NextSeq 2000, and MiSeq™ Systems 600-cycle kit performance comparison.

The availability of 600-cycle kits (Table 1) makes the NextSeq Systems (Figure 3) an excellent choice for labs looking for a powerful, benchtop sequencer for IR-Seq and other NGS applications such as metagenomics and fusion RNA.

Table 1: Options for 600 cycle sequencing runs on NextSeq platforms

NextSeq1000/2000 Flow Cell choices	Catalog #	High-quality data	Clusters/ Read Pairs
P1 Reagents 600-cycle kit	20075294	60 Gb	100 M
P2 Reagents 600-cycle kit	20075295	180 Gb	300 M

The choice of an IR-seq library solution depends on various factors such as input (ie, DNA, bulk RNA, or single-cell RNA), methodology (ie, RACE PCR, or Multiplex PCR), regions sequenced (ie, CDR3 or full length) and the chain type (ie, TCR or BCR). Multiple library prep solutions and protocols for immune repertoire are available offering compelling and differentiated options to choose from (Table 2). We tested these solutions on the NextSeq 2000. Due to insert length above 600bp, many of these options require or benefit from 2 × 300 paired-end sequencing to capture the IR spectrum from the sample.

Table 2: Third-party library prep kit for IR-Seq applications tested on NextSeq 600 cycle kits

Resolution	Method	Provider	Library prep kit name and reference	Targeted receptors	Secondary analysis
Bulk RNA-seq	CDR3/full length	New England Biolabs	NEBNext Immune Sequencing Kit (Human) (E6320S, E620L)	BCR, TCR or BCR + TCR	Open source Presto tools
Bulk RNA-seq	CDR3/full length	QIAGEN	QIAseq Immune Repertoire RNA Library Kit (333705)	TCR	QIAGEN GeneGlobe Analysis
Bulk RNA-seq	CDR3/full length	Takara	SMART-Seq Human TCR (with UMIs) (634780, 634781, 634779)	TCR	Cogent NGS Immune Profiler/Viewer Software
Bulk RNA-seq	CDR3/full length	Takara	SMART-Seq Human BCR (with UMIs) (634777, 634778, 634776)	BCR	Cogent NGS Immune Profiler/Viewer Software
Single cell RNA-seq	CDR3/full length	BD	BD Rhapsody TCR/BCR Multiomic Assay Kit (665828, 665829)	BCR + TCR	BD Rhapsody Sequence Analysis Pipeline

PRIMARY SEQUENCING METRICS

The NextSeq 1000 and NextSeq 2000 deliver high quality PE300 (600-cycle) reads. Initial benchmark testing of the NEBNext Immune Sequencing Kit (New England Biolabs) exceeds Illumina specifications for base and read quality metrics with an overall average %>Q30 of ~91% and PhiX aligned error rate below 0.40% (Figure 4).

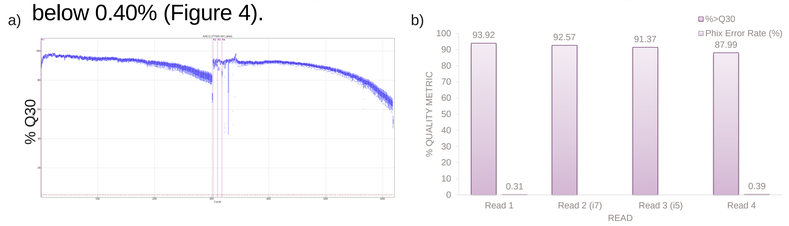

Figure 4: NextSeq 2000 sequencing quality; a) %>Q30 by cycle plot b) average %>Q30 by read plot

SECONDARY IR-SEQ METRICS

Common metrics used in the analysis of B and T cell repertoires include measuring diversity and clonality at various molecular levels. VJ usage (Figure 5a) and clonotype rarefaction plots (Figure 5b) can often show large repertoire level differences between cancer and healthy control patients. In some instances, single-cell immune clustering analysis can further resolve disease signatures through paired chain information (Figure 5c).

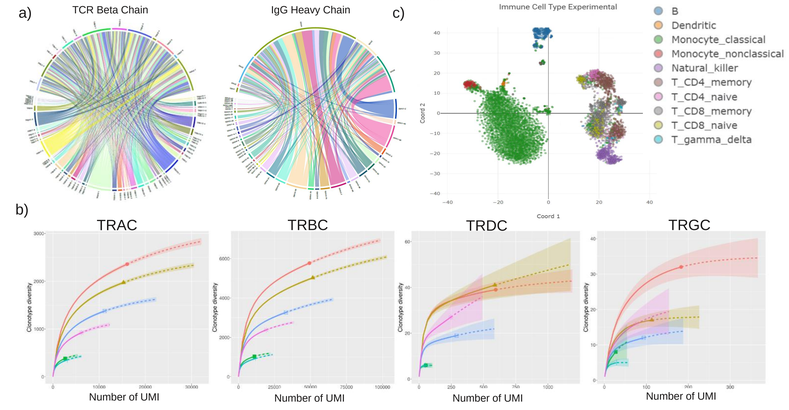

Figure 5: Common secondary metrics from NGS sequencing for IR characterization; a) VJ-usage frequency plots with Takara libraries showing diversity in recombined TCRB (left) and IgG heavy (right) chain genes b) diversity rarefaction plots with QIAGEN libraries are showing impact of RNA input on TCR repertoire in β,α,γ,and δ chains; c) t-SNE plots showing clustering analysis of single immune cell populations and subpopulations based on classical markers with BD libraries.

CONCLUSIONS & REFERENCES

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 quality and quantity of reads necessary for a detailed view of the immune repertoire as well as the capacity for multiplexing of samples.

1. The Antibody Society. Therapeutic monoclonal antibodies approved or in review in the EU or US. 5/6/2023; www.antibodysociety.org/resources/approved-antibodies
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. Laydon DJ, Bangham CR, Asquith B. Estimating T-cell repertoire diversity: limitations of classical estimatorsand a new approach. Philos Trans R Soc Lond B Biol Sci. 2015;370(1675):20140291. doi:10.1098/rstb.2014.0291

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

illumina®