Universal synthetic spike-in controls for accurate adaptive immune receptor profiling

Alex Chenchik, Tianbing Liu, Mikahil Makhanov, Dongfang Hu, Paul Diehl; Cellecta, Inc., Mountain View, CA

Abstract

The results of adaptive immune receptor (AIR) repertoire diversity assays can be affected by various biases from differences in conditions in the RT-PCR and NGS sequencing steps. Spike-in synthetic controls can be used as calibration standards to address these biases. In this study, we synthesized near full-length BCR and TCR constructs that mimic seven different IGH, IGK, IGL, TRB, TRA, TRG and TRD genes.
To test our spike-in controls, three sets of variants at different concentrations were added to the RNA samples before the reverse transcription reaction with Cellecta's DriverMap™ Adaptive Immune Receptor (AIR) Profiling Assay. The DriverMap™ protocol uses reverse gene-specific primers with unique molecular identifiers (UMI), allowing UMI-based correction of amplification biases during data analysis. Calculating UMI-based correction and comparing that with spike-in controls enabled us to differentiate between real and background sequences and estimate the average sequencing error rate at 0.4%-0.8% per base, which is within the reported range of Illumina sequencing. This suggests that our spike-in controls are reliable and may be used as a universal tool to correct AIR protocol biases and calculate error and mutation rates in different AIR profiling assays.

Introduction

- Synthetic TCR/BCR controls work as a calibration standard to measure and compensate for biases caused by the amplification and sequencing of immune receptors.
- Cellecta developed BCR and TCR controls with unique CDR3 regions that mimic the 29 most abundant immune receptor RNA isoforms.
- TCR/BCR RNA calibration standards spiked at different concentrations to experimental RNA samples allowed us to measure sensitivity and linearity in the quantitation of target immune receptor mRNAs.
- Synthetic TCR/BCR control RNAs with nearly full-length V(D)JC structures are designed to be compatible with any RT-PCR-based receptor profiling assays.

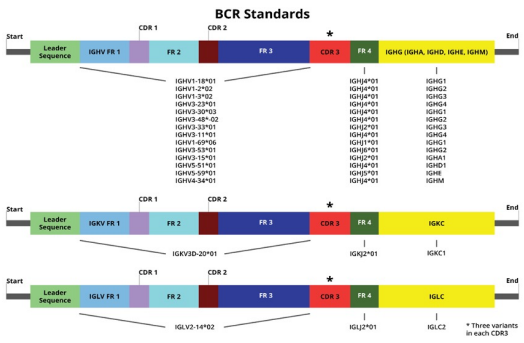


Fig 1: 48 BCR mRNA spike-in constructs represent 10 different IGHs; 1 for each IGHA, IGHD, IGHE, IGHM, IGK and IGL genes.

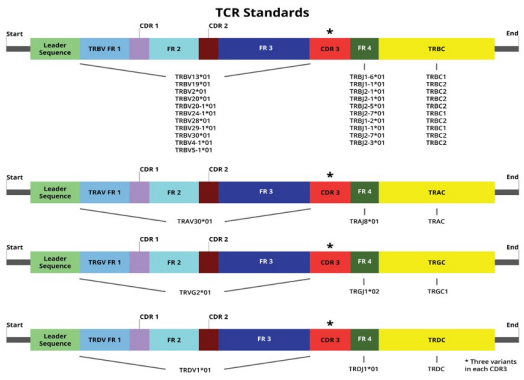
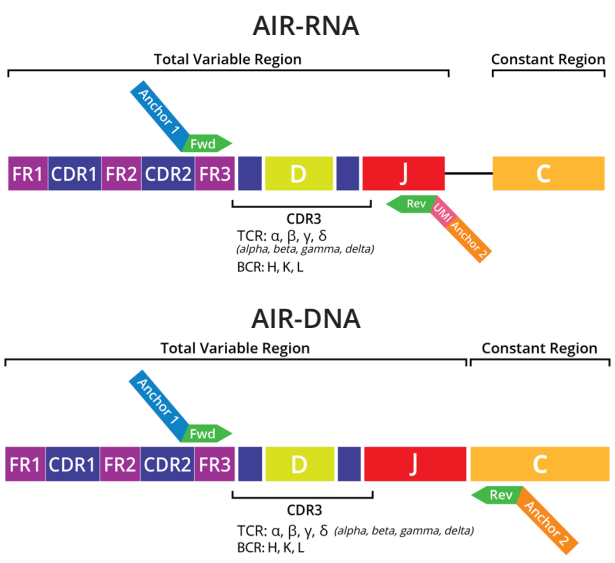


Fig 2: 39 TCR spike-in constructs represent 10 different TRBs, 1 for each of TRA, TRG and TRD genes.

*Each construct has 3 variations in the CDR3 region that differ by 3 nucleotides in a fixed position

DriverMap™ Adaptive Immune Receptor (AIR) Profiling Assays

- **Comprehensive AIR-RNA and DNA repertoire coverage** for all seven TCR and BCR chains in a multiplex RT-PCR reaction (human and mouse).
- **Quantitative clonotype analysis** with UMI for both RNA and DNA templates.
- Integrated with **MiXCR software** (MiLaboratories) package for immune repertoire analysis.

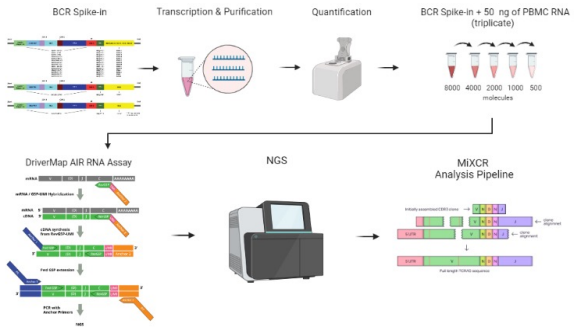


Method

Synthetic TCR/BCR mRNA Control Constructs

- Linear TCR and BCR plasmids were transcribed to make total RNA for each construct.
- A mixture of 48 BCR constructs and 39 TCR constructs was spiked at different concentrations to 50 ng PBMC RNA at a 1:1 ratio.
- DriverMap™ AIR-RNA TCR/BCR assay was run for immune receptor profiling of all TCR (TRA, TRB, TRD, and TRG) and BCR (IGH, IGK, and IGL) chains.
- Pooled library was sequenced on Illumina NextSeq 500, 300-n paired-end kit for 10M reads per sample.
- Data was analyzed using the MiXCR 3.0 software.

Calibration Standard Workflow for RNA Spike-in



Data Analysis Pipeline

Demultiplex samples
bcl2fastq v2.20.0.422 is used for demultiplexing

NGS Data QC
FastQC v0.11.9 is used for Data QC

Clonotype extraction
MiXCR 3.0 and later is used for UMI extraction, indels, mismatches correction, alignment, clonotypes assembly with additional PCR and sequencing errors correction, final clonotypes extraction.

TCR/BCR spike-in Calibration correlation analysis
Customized Python scripts and R scripts were used to analyze TCR/BCR spike in the Calibration dataset.

Results

TCR Calibration Standards
39 TCR constructs with unique CDR3 region

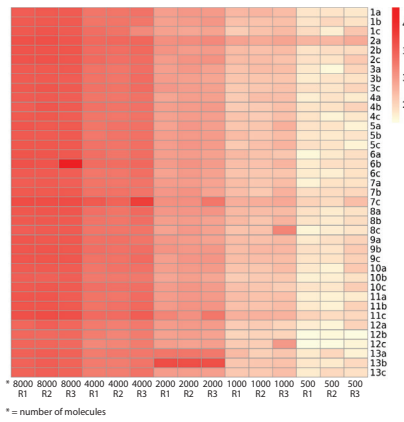


Fig 3: Heatmap of UMI counts in TCR spike-in control RNAs in 13 TCR isoforms (39 sequences) in various concentrations spiked in 50 ng of PBMC RNA.

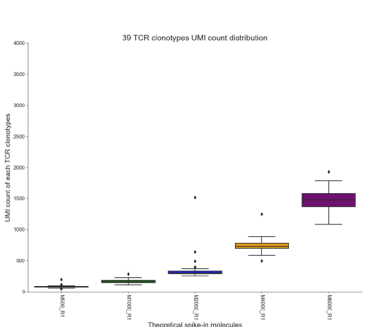


Fig 4: Box plot analysis of TCR spike-in control constructs; results show a linear trend as the number of spike-in molecules increases.

Raw UMI count is used to generate box plot, MiXCR pipeline

Log10 Normalized UMI count =
(UMI count / total UMI count)*500000

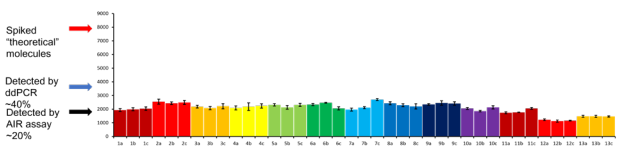


Fig 5: UMI count of 39 TCR control constructs spike-in at 8000 molecules in 50 ng of PBMC RNA

Results (cont.)

BCR Calibration Standards
48 BCR constructs with unique CDR3 region

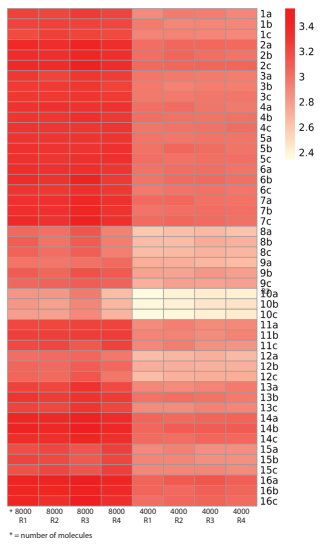


Fig 6: Heatmap of UMI counts in BCR spike-in control RNAs in 16 BCR isoforms (48 sequences) at 8000 and 4000 molecules in 50 ng of PBMC RNA.

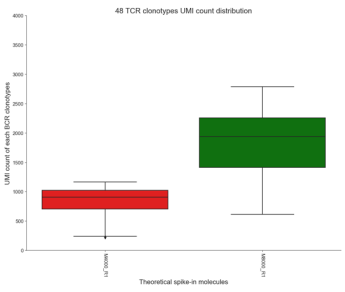


Fig 7: Box plot analysis of BCR spike-in control constructs; results show a linear trend as the number of spike-in molecules increase.

Log10 Normalized UMI count =
(UMI count / total UMI count)*500000

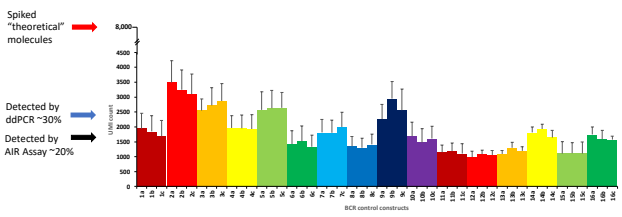


Fig 8: UMI count of 48 BCR control constructs spike-in at 8000 molecules in 50 ng of PBMC RNA

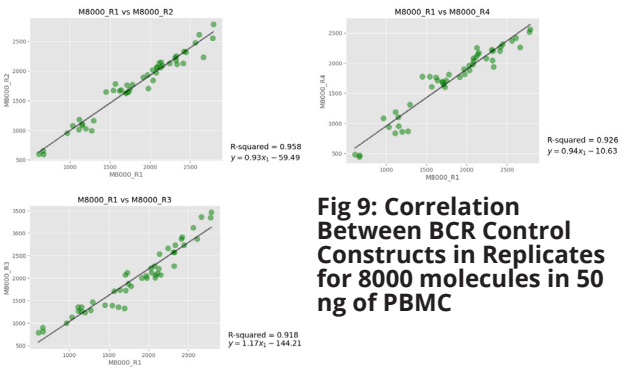


Fig 9: Correlation Between BCR Control Constructs in Replicates for 8000 molecules in 50 ng of PBMC

Discussion

Summary

- Cellecta developed a panel of synthetic, near full-length TCR/BCR RNA and DNA constructs for universal calibration standard application in immune receptor profiling assays.
- TCR/BCR calibration standards can be used for measuring the sensitivity and accuracy (by mutation rate analysis) of multiplex RT-PCR AIR-RNA profiling data using UMI-powered MiXCR analysis in universal PCR-based immune receptor profiling assay
- The design of the calibration standards enables them to be used with any standard AIR-PCR assay.
- Calibration provides a basis to measure primer performance in AIR-PCR assays.

Download Demo Results

Access on GitHub

Access on NCBI:
BioProject ID
PRJNA975292