Development of LymphoTrack® Bioinformatics Methods: Clonality Testing, Somatic Hypermutation and Minimal Residual Disease

Kasey Hutt, Jeff Panganiban, Austin Jacobsen, Jordan Thornes, Jeffrey E Miller, and Michael Klass

Invivoscribe® Technologies, Inc., San Diego, CA, United States

Introduction

- Clonality assessment is an important diagnostic indicator for the malignant transformation of lymphoid cells.
- Minimal Residual Disease (MRD) is a key prognostic indicator for determining efficacy of treatment.
- *IGHV* Somatic Hypermutation (SHM) is a naturally occurring process to increase B-cell receptor diversity, and serves as a positive prognostic factor in Chronic Lymphocytic Leukemia (CLL).
- Next-Generation Sequencing (NGS) assays for measuring these disease markers are becoming the cutting edge of diagnostic methods, but require the development of new bioinformatics tools to confidently analyze and report the results in a simple manner. We have developed the LymphoTrack® bioinformatics tools to analyze *IGH, IGK* and *TRG* genes on the Illumina® MiSeq® and Thermo Fisher Ion PGM® platforms, for the purpose of measuring clonality, *IGHV* SHM and MRD.

Materials and Method

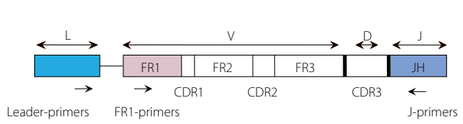
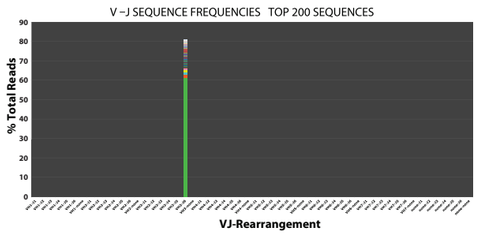


FIGURE 1. PRIMER DESIGN

LymphoTrack® assays have been developed for both the MiSeq® and PGM® NGS platforms, for the *IGH*, *IGK* and *TRG* loci. Each assay employs a separate master mix to amplify the respective upstream V-region (or Leader) with the downstream J-region. Depicted is the *IGH* locus. The MiSeq® assays use 24 indices, allowing for parallel analysis of 22 samples (plus 2 controls). PGM® assays use 12 indices, allowing for parallel analysis of 10 samples (plus 2 controls).

FIGURE 2. BIOINFORMATICS WORKFLOW

LymphoTrack® bioinformatic analysis begins on MiSeq® with stitching of paired reads (PGM® analysis begins at the next step). The resulting sequence file is then filtered for noisy reads (non-specific or fragments that are too short), and reads that do not pass quality are removed. These reads are then aligned to the reference sequence database, and numerous statistics are calculated. SHM rate is calculated from the alignable portion of the V-gene. MRD detection requires the previous identification of a clonal read, and for the user to input the amount of DNA being assayed. MRD results are reported for exact matches, and very similar matches to the clone submitted by the user. A confidence value is assigned to these results that tell the user the reliability of result, which is based on how much DNA was assayed, as well as the read depth achieved for that sample. The confidence is based on theoretical samplings using a binomial distribution.



Rank	Sequence	Length	Raw Count	V-gene	J-gene	% Total Reads	Cumulative %	Mutation Rate to Partial V-Gene (%)	In-Frame (Y/N)	No Stop Codon (Y/N)	V-Coverage
1	GCCTCTGGATTCA…	290	354037	IGHV3-43_01	IGHJ6_02	61.69	61.69	5.24	Y	Y	97.82
2	GCCTCTGGATTCA…	287	8147	IGHV3-43_01	IGHJ6_02	1.42	63.10	5.24	Y	Y	97.82
3	GCCTCTGGATTCA…	290	6798	IGHV3-43_01	IGHJ6_02	1.18	64.29	5.68	Y	Y	97.82
4	GCCTCTGGATTCA…	290	6602	IGHV3-43_01	IGHJ6_02	1.15	65.44	5.68	Y	Y	97.82
5	GCCTCTGGATTCA…	290	4082	IGHV3-43_01	IGHJ6_02	0.71	66.15	5.24	Y	Y	97.82
6	GCCTCTGGATTCA…	290	2364	IGHV3-43_01	IGHJ6_02	0.41	66.56	5.68	Y	Y	97.82
7	GCCTCTGGATTCC…	290	1672	IGHV3-43_01	IGHJ6_02	0.29	66.85	5.68	Y	Y	97.82
8	GCCTCTGGATTCA…	290	1494	IGHV3-43_01	IGHJ6_02	0.26	67.11	5.68	Y	Y	97.82
9	GCCTCTGGATTCA…	290	1361	IGHV3-43_01	IGHJ6_02	0.24	67.35	5.24	Y	Y	97.38
10	GCCTCTGGATTCA…	290	1314	IGHV3-43_01	IGHJ6_02	0.23	67.58	5.68	Y	Y	97.82

FIGURE 3. LYMPHOTRACK® SOFTWARE OUTPUTS

A typical MiSeq® *IGH* FR1 result from a bone marrow sample showing evidence for clonality with one sequence at 62% of total reads and mutation rate of 5%. Top 10 unmerged reads are presented in the table. FASTQ data from either MiSeq® or PGM® can be analyzed by LymphoTrack® Software running on a Windows® PC. The LymphoTrack® Software generates frequency distributions, DNA sequences, V-J assignment and usage, and SHM status. An add-on module allows for the tracking of clonotypes for MRD purposes.

Results

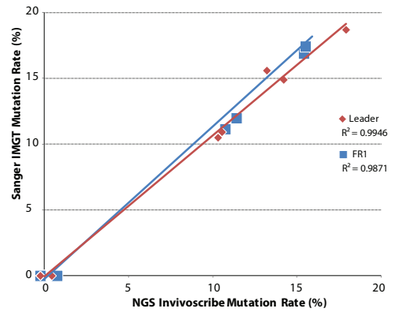


FIGURE 4. SHM CONCORDANCE

LymphoTrack® SHM calculations for *IGHV* FR1 and Leader closely match results generated by Sanger sequencing and IMGT mutation rate calculations.

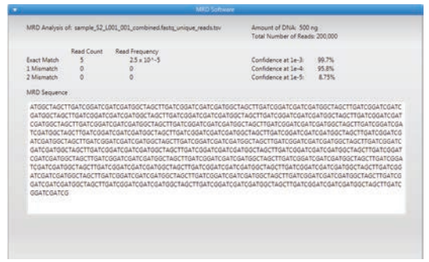


FIGURE 5. MRD OUTPUT EXAMPLE

Example output for MRD module. Results include the read count and frequency, as well as confidence metrics for various levels of detection. Additional outputs will include sequence files for all reads.

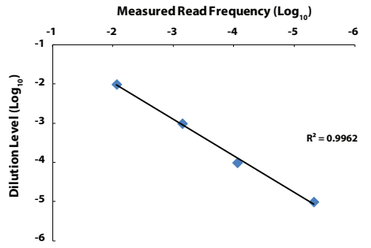


FIGURE 6. MRD DILUTION CONCORDANCE

Dilution experiment shows strong linear concordance between dilution level and measured read frequency.

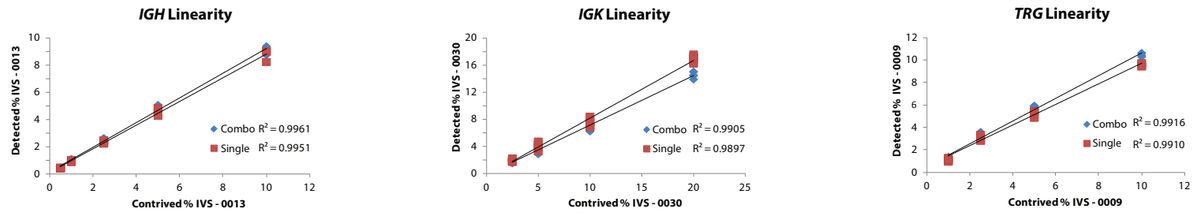


FIGURE 7. MULTIPLEXING CAPABILITY

LymphoTrack® Software is able to analyze samples mixed with all different target loci (*IGH*, *IGK* and *TRG*) even when sharing barcodes without a significant change in signal.

Conclusions

- LymphoTrack® bioinformatics analysis accomodates a wide variety of important diagnostic and prognostic indicators. The straightforward presentation of results ensures an equally uncomplicated determination of clonality, SHM, and MRD.
- LymphoTrack® bioinformatics analysis also provides the resources necessary to delve deeper into the data by providing full sequence information and V-J assignment, allowing for an unprecedented level of granularity previously unattainable using capillary electrophoresis data.