High throughout sequencing and IMGT/HighV-QUEST analysis of T cell receptor repertoire

Eric Gowans (eric.gowans@adelaide.edu.au)
University of Adelaide, Australia
Marie-Paule Lefranc
Université Montpellier 2, France
Shuo Li
Véronique Giudicelli
Université Montpellier 2, France

Method Article

Keywords: T cell receptor, IMGT/HighV-QUEST

Posted Date: September 5th, 2013

DOI: https://doi.org/10.1038/protex.2013.072

Abstract

We describe a workflow combing 5'RACE PCR, 454 sequencing and IMGT/HighV-QUEST for profiling human T cell receptor repertoire.

Procedure

1. T cells were sorted from PBMC by flow cytometry. 2. RNAwas immediately extracted from sorted cells using RNeasy minikit \(Qiagen). 3. A 5' rapid amplification of cDNA ends \(5'RACE) PCR was conducted using the SMARTer™ RACE cDNA Amplification Kit \(Clontech Laboratories) according to the manufacturer's instructions. We use proximally 200 ngRNA for each library. The extension time for the first-strand cDNA synthesis was 90 min at 42°C followed by 15 min inactivation at 70°C. 4. The first-round PCR was achieved using Phusion Hot-StartDNA Polymerase \(Finnzymes), a template switching oligonucleotide \(TSO, supplied in the above SMARTer™ RACE cDNA Amplification Kit), a universal primer mix \(supplied in the above SMARTer™ RACE cDNA Amplification Kit), along with the TRBC gene-specific reverse primer, 5'ttctgatggctcaaacac3' \(codon positions 11-6, IMGT unique numbering) which aligns to both TRBC1 and TRBC2 \(IMGT Repertoire, http://www.imgt.org). The cycling conditions were: 30 sec denaturation at 98°C, 26 cycles of 10 sec at 98°C, 10 sec at 55°C, and 20 sec at 72°C, plus a final extension for 5 min at 72°C. 5. The reaction products were purified using QIAquick columns \(Qiagen). The purified DNA fragment was loaded on a 1.5% low melting temperature agarose gel, and a band corresponding to a 500–650 bp product was excised and purified using the QIAquick Gel Extraction Kit \(Qiagen). 6. A second-round PCR was performed on a fraction of the first-round reaction. This step incorporated Roche forward and reverse linker primers to enable the sequencing and barcodes to distinguish different samples \(454 Sequencing Technical bulletin TCB N°013-2009, August 2009). 7. The product of the second-round PCR was purified and quantified using PicoGreen reagent \(Invitrogen). 8. Sequencing was performed using the 454 Genome SequencerFLX \(GSFLX) Titanium \(Roche). 9. Initial data processing was performed using the manufacturer's software, which included the removal of low quality and erroneous sequences as determined by the standard filters of the Roche amplicon signal-processing pipeline. Sequences were assigned to samples based on incorporated barcodes, and read orientation was determined by the presence or absence of the sequence corresponding to the template switching oligonucleotide \(TSO) used in theSMARTer™ RACE. Sequence segments corresponding to the adapters, barcodes and TSO were removed during this process. 10. The sequences obtained from step 9 were submitted online to IMGT/HighV-QUEST. The full capacity of IMGT/HighV-QUEST includes analysis of V-J and V-D-J rearranged sequences \(up to 150,000 per job) and statistical analysis \(on results of up to 450,000 sequences) \(http://www.imgt.org, version July 2012). The IMGT/HighV-QUEST submission page allows users to submit a file containing up to 150,000 sequences and to select options \(equivalent to those of IMGT/V-QUEST) for the results display. 11. The results are provided in a downloadable main folder with eleven files in CSV format \(results equivalent to those of the Excel file from IMGT/V-QUEST online), and one folder with the individual files \(up to 150,000) of all the sequence results. For each analysed sequence, the results in those individual files are identical to those that could be obtained from

IMGT/V-QUEST online \(in display option ‘Text’ of ‘Detailed view’). Text and CSV formats facilitate statistical studies for further interpretation and information extraction. Prior to IMGT/HighV-QUEST analysis, the users can evaluate the quality of their sequences by checking the results obtained with IMGT/V-QUEST on a few sequences. 12. In a second online step, the users can submit the results of one or several jobs \(up to 450,000 results) for statistical analysis. The IMGT/HighV-QUEST ‘Summary’ table of the statistical analysis provides information in Results categories that are either filtered in \(‘1 copy’, ‘More than 1’) or filtered out \(‘Warnings’, ‘Unknown functionality’, ‘No results’). The number of sequences in the different categories provides the users with an immediate indication of data reliability. 13. The 5’ and 3’ reads were pooled prior to genotype and haplotypes identification. This is to overcome the limitation of 454 sequencing, which does not provide genuine “bi-directional” sequences. Indeed the 5’ reads and 3’ reads are generated independently in separate wells. The genotype and haplotypes were deduced from the IMGT/HighV-QUEST statistical analysis performed on the pooled data on the results category ‘1 copy’ ‘single allele’ \(for V and J). Referenced: 1. Alamyar, E., Giudicelli, V., Li, S., Duroux, P. & Lefranc, M.-P. IMGT/HighV-QUEST: the IMGT® web portal for immunoglobulin \(IG) or antibody and T cell receptor \(TR) analysis from NGS high throughput and deep sequencing. Immunome Res 8, 26 \(2012). 2. Alamyar, E., Duroux, P., Lefranc, M.-P. & Giudicelli, V. IMGT® tools for the nucleotide analysis of immunoglobulin \(IG) and T cell receptor \(TR) V-\(D)-J repertoires, polymorphisms, and IG mutations:IMGT/V-QUEST and IMGT/HighV-QUEST for NGS. Methods Mol Biol 882, 569-604 \(2012). 3. Freeman, J.D., Warren, R.L., Webb, J.R., Nelson, B.H. & Holt, R.A. Profiling the T-cell receptor beta-chain repertoire by massively parallel sequencing. Genome Res 19, 1817-1824 \(2009).