

Immunological Genome

Project

Immgen.org

ImmGen Poly-A+ RNA sequencing (full depth) 11-cell set

Full depth directional RNA sequencing was performed on the core ImmGen "11 -cell set" populations to generate reference datasets for the

Tissues from 5 week-old C57BL/6J (Jackson Laboratory) males, double-sorted by flow cytometry, per ImmGen cell preparation SOP.

From the Spleen (mechanical disruption)

1. Total CD4 (TCRB+CD4+)
2. Total CD8 (TCRB+CD8+)
3. Total CD19+ B (CD19+IgM+)
4. Treg (TCRB+CD4+CD25+)
5. NK (NK1.1+TCRB-)
6. TCRgd (TCRgd+TCRB-)
7. DC (CD11c+MHCII+ FLT3+)
8. Neutrophils (LY66+)
9. NKT (TCRBint NK1.1int)
10. NKT (CD1d Tetramer Sorted, TCRB int)

From the Peritoneal Cavity lavage (10 ml PBS)

11. B1ab cells (CD19+CD43+CD5+)
12. PC Macrophages (F4/80+ICAM2+)

RNA was prepared from 100,000 cells (final sort collected directly in Trizol). Total RNA was extracted from each sample following ImmGen RNA extraction SOP.

PolyA+ RNA was isolated by bead capture (PrepX PolyA mRNA Isolation Kit, Wafergen Biosystems), following the manufacturer's instructions.

Then mRNA libraries were prepared with the PrepX RNA-Seq Library Preparation Kit (Integenex). For first strand synthesis, mRNA was fragmented and ligated to 3' and 5' adapters before reverse transcription. Sample libraries then underwent bead clean-up, 2nd strand synthesis and PCR amplification before final sequencing. Libraries were analyzed on an Illumina HiSeq2500 sequencer.

fastq files are generated by Illumina's CASAVA

(<http://www.illumina.com/informatics/sequencing-microarray-data-analysis/casava.ilmn>)

and demultiplexed. For data analysis, sequence quality for each sample was checked by fastqc (ver. 1.10.1, <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>). No adapter sequences or primer sequences were detected. 10th percent of all the base quality score were greater than 28. Alignment to the mouse genome (UCSC mm10 genome) was performed with Tophat (ver. 2.0.10, <http://ccb.jhu.edu/software/tophat/index.shtml>) on a high-performance compute cluster at Harvard Medical School. The alignment result files in bam file format were later converted to bedgraph format to display on UCSC genome browser (

http://genome.ucsc.edu/cgi-bin/hgTracks?hgS_doOtherUser=submit&hgS_otherUserName=hp65&hgS_otherUserSessionName=ImmGen_11). The gene-level quantification and normalization were computed by htseq-count (HTSeq package 0.6.1, <http://www-huber.embl.de/users/anders/HTSeq/doc/count.html>)

and DESeq package from BioConductor
(<http://bioconductor.org/packages/release/bioc/html/DESeq.html>).

The read depth and quality parameters for the 11-cell-set sequencing are shown in Table 1.

Population	Short name	Purity	Total mapped reads
Spleen Neutrophil	GN.SP#1	95.4%	17,227,773
Spleen Dendritic Cells	DC.SP#1	95.7%	15,682,458
Peritoneal Cavity B1ab Cells	B1ab.PC#1	95.8%	21,34,7470
Spleen T regulatory cells	Treg.Sp#1	95.2%	16,565,681
Spleen NK Cell	NK.SP#1	89.7%	15,887,314
Spleen CD4 T Cell	T4.Sp#1	95.8%	16,230,633
Spleen CD8 T cell	CD8 T Cell	95.4%	25,324,688
Spleen CD19+ B cell	B Cell	95.6%	18,947,445
Spleen TRCgd	gdT Cell	88.8%	8,026,274
Spleen NKT	NKT.SP2	92.6%	23,124,482
Peritoneal Macrophage	MF.PC	93.0%	8,196,538