HHS Public Access

Author manuscript

J Bioinform Comput Biol. Author manuscript; available in PMC 2016 April 01.

Published in final edited form as:

J Bioinform Comput Biol. 2015 April ; 13(2): 1550010. doi:10.1142/S0219720015500109.

Analytical workflow profiling gene expression in murine macrophages

Scott E. Nixon[*], **Dianelys González-Peña**[†], **Marcus A. Lawson**[‡], **Robert H. McCusker**[†], **Alvaro G. Hernandez**[§], **Jason C. O'Connor**[¶], **Robert Dantzer**[||], **Keith W. Kelley**[†], and **Sandra L. Rodriguez-Zas**[†,**,††]

Sandra L. Rodriguez-Zas: rodrgzzs@illinois.edu

[*]Illinois Informatics Institute, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

[†]Department of Animal Sciences, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

[‡]Division of Nutritional Sciences, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

[§]Roy J. Carver Biotechnology Center, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

[¶]Department of Pharmacology, University of Texas Health Science Center at San Antonio, San Antonio, TX 78229, USA

[||]Department of Symptom Research, University of Texas M. D. Anderson Cancer Center, Houston, TX 77030, USA

[**]Department of Statistics, University of Illinois Urbana-Champaign, Urbana, IL 61801, USA

[††]Institute for Genomic Biology, University of Illinois Urbana-Champaign, Urbana, IL 61801, USA

Abstract

Comprehensive and simultaneous analysis of all genes in a biological sample is a capability of RNA-Seq technology. Analysis of the entire transcriptome benefits from summarization of genes at the functional level. As a cellular response of interest not previously explored with RNA-Seq, peritoneal macrophages from mice under two conditions (control and immunologically challenged) were analyzed for gene expression differences. Quantification of individual transcripts modeled RNA-Seq read distribution and uncertainty (using a Beta Negative Binomial distribution), then tested for differential transcript expression (False Discovery Rate-adjusted p-value < 0.05). Enrichment of functional categories utilized the list of differentially expressed genes. A total of 2079 differentially expressed transcripts representing 1884 genes were detected. Enrichment of 92 categories from Gene Ontology Biological Processes and Molecular Functions, and KEGG pathways were grouped into 6 clusters. Clusters included defense and inflammatory response (Enrichment Score = 11.24) and ribosomal activity (Enrichment Score = 17.89). Our work provides a context to the fine detail of individual gene expression differences in murine peritoneal macrophages during immunological challenge with high throughput RNA-Seq.

Correspondence to: Sandra L. Rodriguez-Zas, rodrgzzs@illinois.edu.

Keywords

RNA-Seq; functional analysis; transcriptome; macrophage

1. Introduction

Identification and analysis of an individual gene may offer limited insights. While genes serve as one of the smallest units by which biological change can be measured, critical information comes from considering the sum of their individual effects. Expanding the "snapshot view" available for differential expression motivates a drive towards the enlargement of analyses from single gene studies with quantitative real-time PCR to microarrays, and more recently RNA-Seq.[1]

The range of tools available for RNA-Seq analysis, as well as the tools themselves, undergoes a rapid pace of modification. These changes demand a thorough understanding of how the tools operate to choose appropriate settings for a particular experiment. Without a singular accepted method or settings to address all applications, transcriptomics relies upon the validation of data quality and controls.[2] TopHat, Cufflinks, and Cuffdiff comprise a set of tools for analyzing RNA-Seq datasets.[3] These tools have gained popularity for the capability to handle intron-spanning reads, and options to address various biological- and technical-biases that are of concern during analysis.[4]

Mapping RNA reads to an annotated genome is one of the popular and well-established methods for differential expression testing between treatments in model organisms.[5] With the potential to detect thousands of differentially expressed genes, organizing these differences into more interpretable groups becomes the purpose of downstream tools. One possibility that is explored here involves grouping the gene information into groups based upon their functional actions, a form of gene set enrichment.[6] This study examines the capability of a RNA-Seq-based workflow to evaluate transcriptomic changes. Efficient identification of differentially expressed genes and the functions they impact elucidates their modification of the biological state between treatments. The novelty of this experiment is in the application of RNA-Seq and the associated algorithms to a particular biological model, the analysis of peritoneal macrophages from Bacille Calmette–Guérin (BCG)-challenged mice compared to those receiving a saline control.[7] This challenge has been associated with substantial changes in sickness and depression-like indicators.[8] The characterization of the transcriptome during immunological resolution and behavioral transition seven days after initial challenge is of interest. RNA-Seq has yet to be applied to characterize the transcriptome at this time point.[9–11] The application of RNA-Seq and downstream methods to analyze changes in transcriptomics in this model has not been reported, providing a new level of capability in constructing an inflammation-induced immunological response profile. Transcript profiles were further studied and interpreted using functional enrichment analyses to uncover categories that may be over-represented among particular profiles.

2. Materials and Methods

RNA-Seq technology was used to study changes in gene expression in macrophages taken from mice following a previously established immune-challenge model.[10] Male adult (~ 22 weeks of age; n = 6/group) C57BL/6J mice were injected into the peritoneum with TICE strain BCG (Organon USA Inc., USA) or equal volume (10 mg) physiological saline (Control). Utilizing the same inbred strain as used for the Mouse Genome Project minimizes genetic variations that could hinder mapping.[12] RNA was isolated from macrophages collected from the peritoneal cavity seven days post-challenge.[13,14] This timing of collection was selected to capture transcriptome changes during a period of immunological and behavioral transitions.[8,11]

The workflow of RNA-Seq data analysis is presented in Fig. 1. Transcriptomic analysis with RNA-Seq involves producing libraries of reads that represent gene transcripts from the samples for quantitative comparison. Individual mouse RNA-Seq libraries were sequenced using Illumina HiSeq2000 (Illumina, San Diego, CA) to produce paired-end 100-bp reads, summarized as “left” and “right” reads. One library of reads per biological sample was examined for sequencing errors prior to mapping to genome and transcriptome features. Quality control of sequence reads used FastQC (Fig. 2).[15] Quality was determined by the reported score at each base position (> 30), a Qphred quality value which is the negative logarithmic transformation of the estimated probability of error (Eq. (1)).[16]

$$Q_{\mathrm{PHRED}} = -10 \times \log_{10}(P_e). \quad (1)$$

Reads were mapped to the mouse genome (GRCm38) and assembled using TopHat2 (TopHat v2.0.9) and Cufflinks and analyzed using Cuffmerge, and Cuffdiff 2 (v2.1.1, Fig. 1).[3] TopHat2 maps reads via the use of Bowtie2, the core read-alignment program, while TopHat2 deals with splicing concerns from mapping intron-spanning RNA reads to a DNA genome.[3] Due to the computational scale of mapping millions of reads to large genomes, Bowtie2 implements Burrows–Wheeler transformation to efficiently scan the genome during mapping.[17] TopHat2 was chosen for its two-step method to deal with spliced alignments and preferential alignment of reads onto real genes from an annotation.[18]

Reads were assembled based upon mapping information into gene transcripts, with transcripts quantified by condition for differential comparison as elaborated in Ref. 3. The Cufflinks program (http://cufflinks.cbcb.umd.edu/) takes the mapping information from TopHat2, and assembles the reads back into the biologically relevant transcripts that would have produced them. Cufflinks offers optional assembly methods that correct for biological and technical biases, including biases in Illumina’s read-creation process.[19] Options to correct for fragment bias during transcription priming with random hexamers and estimation of appropriate counting for those reads that can map to multiple sites were used.[20,21] Upper Quartile normalization was enabled for its superior performance compared to the default Total Count method available in Cufflinks.[22]

Cuffdiff 2 (referred to from here on simply as “Cuffdiff”) performs differential expression testing between conditions by checking if each gene follows a beta negative binomial

distribution. The beta negative binomial distribution can account for potential overdispersion between groups or uncertainty in read counts that may otherwise be ignored by simpler models.[4] Before any testing for significance, all loci in the genome first needed a minimum number of fragment alignments (10 fragments; Test Status "OK"). Genes within a locus could be analyzed for significance after this minimum alignment (MA) within Cuffdiff was satisfied. Of those genes in locations with > 10 fragment alignments, a list of genes exhibiting significant differential expression between conditions (False Discovery Rate or FDR-adjusted p-value < 0.05) was obtained. The genes were named based upon annotation available from the UCSC database (see Ref. 23, www.genome.ucsc.edu).

Two complementary approaches were used to identify functional categories among transcript profiles. Enrichment based on the hypergeometric test applied to a list of differentially abundant transcript isoforms and gene set enrichment analysis (GSEA) of all transcript isoforms based on the Kolmogorov–Smirnov statistics were evaluated.[24] Gene Ontological (GO, see Ref. 25, www.geneontology.org) terms related to Biological Process (BP) and Molecular Function (MF) were tested, along with the Kyoto Encyclopedia of Genes and Genomes (KEGG)-Pathway database (see Ref. 26, http://www.genome.jp/kegg). For the hypergeometric test, functional category enrichment and functional annotation clustering were performed in the Database for Annotation, Visualization, and Integrated Discovery (DAVID).[6] Specifically the GO FAT categories within DAVID were tested, a filter of GO categories to minimize repetition of general categories and to focus on more specific term identification. Individual categories in DAVID are deemed enriched by using a one-tailed jackknifed Fisher exact test, the EASE score.[27] The downstream functional annotation clustering of these categories used Enrichment Score (ES), calculated as the −log scale geometric mean of the EASE scores of member categories.[28]

For the purposes of clustering, DAVID by default considers categories individually by their EASE score (EASE ≤ 0.1) without concern for experiment-wide false-detection.[27] To avoid errors related to multiple tests, categories were only considered enriched if they were significant at FDR-adjusted EASE score based p-value < 0.1.[29] Cluster ES were re-calculated to reflect the remaining member categories (ES > 4). The GSEA methodology was implemented using the software package GSEA-P and enrichment was tested against the functional categories present in the Molecular Signature Database (MSigDB).[24] The recommended GSEA FDR-adjusted p-value < 0.25 threshold was used in agreement with the statistical testing implemented.[24] Categories consistent between the hypergeometric and GSEA approaches are reported and discussed. These results are robust to differences in assumptions and methodologies between the approaches.

3. Results

The quality control was evaluated for every sample. No evidence of low quality reads was observed within the samples, with quality scores greater than 30 across the entire length of the reads. Quality scores were similarly high across both Control and BCG groups (Fig. 2). Scores ranged between 30 and 40, indicating accuracies between 99.9% and 99.99% for the bases at those positions. Based upon the observed quality of the sample data as well as the read filtering internal to TopHat2, trimming was not needed.

The RNA-Seq reads produced 54 ± 8.5 million reads and 64 ± 6 million paired-end reads of 100 bp in length per sample for Control and BCG, respectively. On average 91% of total reads was mapped to the genome for both Control and BCG. The percentage of reads per sample that successfully mapped to the genome ranged from 74% to 95%, the percentage of reads that produced aligning pairs was also in the same range (Table 1).

Following evaluation of read quality, their assembly into transcripts produced over 60,000 transcripts among all samples. Prior to differential testing between the groups, these transcripts were filtered based upon sufficient alignment coverage and experiment-wide significance cut-offs. The number of differentially expressed transcripts between Control and BCG groups was 2079 (Table 2; 1373: FDR p-value < 0.01; 706: 0.01 < FDR p-value < 0.05), representing 1884 genes.

Among the differentially expressed genes, 802 were under-expressed in BCG vs. Control, indicating similar quantity of up- and down-regulated genes post-challenge. However, there was a predominance of genes overexpressed in BCG relative to Control among the most significant profiles. The most significantly differentially expressed genes (FDR p-value < 0.01) are listed separately for those over- (Table 3) and underexpressed (Table 4) in BCG vs. Control, together with supporting references when previously associated with macrophage populations and their immunological response profile.

Functional categorization of the gene list resulted in 92 significantly enriched terms (BP: 69 terms; MF: 20 terms; KEGG: 3 terms; FDR p-value < 0.1; not listed). Clustering the enriched terms further reduced the list to 6 highly enriched clusters (ES > 4), listed by score in Table 5. Clusters were dominated by GO BP terms as they were the majority of the significantly enriched term list, with terms in the clusters underscoring the activation and regulation of the immune system following challenge. These clusters accounted for 24 of the significantly enriched terms.

4. Discussion

Quality control of the input reads is an important step to successful downstream mapping. Once the reads were determined to be of high quality, the filtering controls implemented by TopHat2 prior to mapping made additional trimming of the reads unnecessary (Table 1).[54] Percentages of mapped reads were similar to those reported in previous high-stringency methods, and approached the percentages seen when previously tested on simulated error-free data.[18,55] The mapping capability of aligners like TopHat2 is dependent upon the genome and annotation, meaning unmapped reads may include those associated with transcripts not yet represented in the annotation. Findings from these RNA-Seq confirmed several results from previous studies that used similar models and quantitative real-time PCR or microarray technologies and uncovered additional profiles and enriched categories. This study centered on one type of peripheral macrophage, collected at one time point and using a specific collection method on macrophage activation status. A longitudinal study of additional macrophage populations using alternative collection methods is necessary for extrapolation of our findings to wider conditions.

The workflow described here effectively identifies genes that are differentially expressed during an immunological challenge and clusters these results based upon functionality. Significantly differentially expressed genes illustrated the extended expression response after BCG-challenge. Among the overexpressed genes, the most overexpressed gene S100a9 works as a heterodimer with S100a8, also found to be in the overexpressed list (Table 3). As both are associated with inflammatory events and are inducible in mature macrophages, their presence after BCG-challenge is expected.[56] The overexpression of CCL5 and CXCL10 (Table 3) was also unsurprising, considering the inhibitory action of IL-10 upon both, and that reduced IL-10 levels were associated with increased resistance to intracellular pathogens[49] such as BCG (Il10, Table 4). The overexpression of Arg1 in the BCG group is consistent with previous work studying the effect of this challenge in macrophages.[57] Along with the most underexpressed gene in BCG compared to Control, Retnla, these indicate underexpression of Th2-associated genes due to the classic Th1-response to BCG.[45,58] It is interesting to find Mt1 and Il10 together in the underexpressed category. Although studies were previously performed in T cells, Mt1/Mt2-deficient mice were found to produce increased levels of Il10 following an immune challenge with anti-CD3/CD28.[59] Still, the role of metallothionein genes during immune challenges and inflammation are not fully elucidated, and low expression of Mt1 supports the proinflammatory nature of the response at the measured time-point.[60] Direct association in the literature between macrophages and the underexpressed gene Ptprcap was less clear, although it has been found in the monocyte precursors to macrophages.[61] However, Ptprcap is known as a CD45-associate, regulating the interaction of CD45 with other proteins. As CD45 regulates apoptosis, this may explain the relationship to immune-challenge.[62] Although the number of differentially expressed genes were similarly split between overexpressed and underexpressed in the BCG relative to the control group, a more stringent significance cutoff found a predominance of genes overexpressed in the BCG group. These results are consistent with other reports of overexpression in the microglia of genes associated with inflammation response in response to an inflammatory challenge.[63,64]

Table 5 summarizes the enriched functional categories consistently detected by the hypergeometric test and GSEA approaches. Enrichment analysis highlighted the biological response of macrophages to an immunological challenge (inflammation-based defense responses; clusters 2 and 3 in Table 5). Clustering was effective at identifying cytokine and chemokine activity in immune cells that are typically associated with activation of macrophages.[65] Categories previously associated with similar immune challenges[66] were clustered to better clarify the transcriptomic differences between experimental groups. Several genes that were overexpressed in BCG relative to Control (Table 3) are affiliated to regulation of locomotion (Table 5) including Xcl1, Cxcl13, Cxcr2, Cxcl10, and Ccl5. These associations could be related to the typical amelioration of sickness behaviors and higher activity observed in mice seven days post-challenge.[8,10] A ribosomal cluster (cluster 1) dominating the list is expected, as protein regulation is at the core of immunological response.[65] The gene lists and resulting clusters from RNA-Seq technology allows for analysis based upon the shared and unique genes. In future studies, this response profile of immunologically challenged peritoneal macrophages can be compared to similar constructed

profiles of other cell populations or challenges to identify profile characteristics unique to each combination.

Acknowledgments

The support of NIH grant numbers: R21 MH 096030, R01 MH 090127, R01 SUB UT 00000712, R01 MH083767, and USDA NIFA grant numbers ILLU-538-632 and ILLU-538-909 are greatly appreciated.

References

1. Oshlack A, Robinson MD, Young MD. From RNA-Seq reads to differential expression results. Genome Biol. 2010; 11(12):220.10.1186/gb-2010-11-12-220 [PubMed: 21176179]
2. Van Verk MC, Hickman R, Pieterse CM, Van Wees SC. RNA-Seq: Revelation of the messengers. Trends Plant Sci. 2013; 18(4):175–179.10.1016/j.tplants.2013.02.001 [PubMed: 23481128]
3. Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, Pimentel H, Salzberg SL, Rinn JL, Pachter L. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat Protoc. 2012; 7(3):562–578.10.1038/nprot.2012.016 [PubMed: 22383036]
4. Trapnell C, Hendrickson DG, Sauvageau M, Goff L, Rinn RL, Pachter L. Differential analysis of gene regulation at transcript resolution with RNA-seq. Nat Biotechnol. 2013; 31(1):46–53.10.1038/nbt.2450 [PubMed: 23222703]
5. Garber M, Grabherr MG, Guttman M, Trapnell C. Computational methods for transcriptome annotation and quantification using RNA-Seq. Nat Methods. 2011; 8(6):469–477.10.1038/nmeth.1613 [PubMed: 21623353]
6. Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc. 2009; 4(1):44–57.10.1038/nprot.2008.211 [PubMed: 19131956]
7. Moreau M, Lestage J, Verrier D, Mormède C, Kelley KW, Dantzer R, Castanon N. Bacille Calmette-Guèrin inoculation induces chronic activation of peripheral and brain indoleamine 2,3-dioxygenase in mice. J Infect Dis. 2005; 192(3):537–544.10.1086/431603 [PubMed: 15995970]
8. Rodriguez-Zas SL, Nixon SE, Lawson MA, Mccusker RH, Southey BR, O'Connor JC, Dantzer R, Kelley KW. Advancing the understanding of behaviors associated with Bacille Calmette Guèrin infection using multivariate analysis. Brain Behav Immun. 201410.1016/j.bbi.2014.09.018
9. Moreau M, Andrè C, O'Connor JC, Dumich SA, Woods JA, Kelley KW, Dantzer R, Lestage J, Castanon N. Inoculation of Bacillus Calmette-Guèrin to mice induces an acute episode of sickness behavior followed by chronic depressive-like behavior. Brain Behav Immun. 2008; 22(7):1087–1095.10.1016/j.bbi.2008.04.001 [PubMed: 18479887]
10. O'Connor JC, Lawson MA, Andrè C, Briley EM, Szegedi SS, Lestage J, Castanon N, Herkenham M, Dantzer R, Kelley KW. Induction of IDO by Bacille Calmette-Guèrin is responsible for development of murine depressive-like behavior. J Immunol. 2009; 182(5):3202–3212.10.4049/jimmunol.0802722 [PubMed: 19234218]
11. Platt B, Schulenberg J, Klee N, Nizami M, Clark JA. A depressive phenotype induced by Bacille Calmette Guérin in 'susceptible' animals: Sensitivity to antidepressants. Psychopharmacology. 2013; 226(3):501–513.10.1007/s00213-012-2923-6 [PubMed: 23179968]
12. Levy S, Sutton G, Ng PC, Feuk L, Halpern AL, Walenz BP, Axelrod N, Huang J, Kirkness EF, Denisov G, Lin Y, MacDonald JR, Pang AWC, Shago M, Stockwell TB, Tsiamouri A, Bafna V, Bansal V, Kravitz SA, Busam DA, Beeson KY, McIntosh TC, Remington KA, Abril JF, Gill J, Borman J, Rogers YH, Frazier ME, Scherer SW, Strausberg RL, Ventner JC. The diploid genome sequence of an individual human. PLoS Biol. 2007; 5(10):e254.10.1371/journal.pbio.0050254 [PubMed: 17803354]
13. Reno C, Marchuk L, Sciore P, Frank CB, Hart DA. Rapid isolation of total RNA from small samples of hypocellular, dense connective tissues. BioTechniques. 1997; 22:1082–1086. [PubMed: 9187757]

14. Zhang X, Goncalves R, Mosser DM. The isolation and characterization of murine macrophages. Curr Protoc Immunol. 2008; 83(14.1):1–14.10.1002/0471142735.im1401s83 [PubMed: 19016445]
15. The Babraham Institute. FastQC. 2013. Available at http://www.bioinformatics.bab-raham.ac.uk/projects/fastqc/
16. Cock PJ, Fields CJ, Goto N, Heuer ML, Rice PM. The Sanger FASTQ file format for sequences with quality scores, and the Solexa/Illumina FASTQ variants. Nucleic Acids Res. 2010; 38(6): 1767–1771.10.1093/nar/gkp1137 [PubMed: 20015970]
17. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie2. Nat Methods. 2012; 9(4): 357–359.10.1038/nmeth.1923 [PubMed: 22388286]
18. Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. TopHat2: Accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol. 2013; 14(4):R36.10.1186/gb-2013-14-4-r36 [PubMed: 23618408]
19. Hansen KD, Brenner SE, Dudoit S. Biases in Illumina transcriptome sequencing caused by random hexamer priming. Nucleic Acids Res. 2010; 38(12):e131.10.1093/nar/gkq224 [PubMed: 20395217]
20. Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B. Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat Methods. 2008; 5:621–628.10.1038/nmeth.1226 [PubMed: 18516045]
21. Roberts A, Trapnell C, Donaghey J, Rinn JL, Pachter L. Improving RNA-Seq expression estimates by correcting for fragment bias. Genome Biol. 2011; 12(3):R22.10.1186/gb-2011-12-3-r22 [PubMed: 21410973]
22. Dillies MA, Rau A, Aubert J, Hennequet-Antier C, Jeanmougin M, Servant N, Keime C, Marot G, Castel D, Estelle J, Guernec G, Jagla B, Jouneau L, Laloë D, Le Gall C, Schaëffer B, Le Crom S, Guedj M, Jaffrèzic F. A comprehensive evaluation of normalization methods for Illumina high-throughput RNA sequencing data analysis. Brief Bioinform. 2013; 14(6):671–683.10.1093/bib/bbs046 [PubMed: 22988256]
23. Karolchik D, Barber GP, Casper J, Clawson H, Cline MS, Diekhans M, Dreszer TR, Fujita PA, Guruvadoo L, Haeussler M, Harte RA, Heitner S, Hinrichs AS, Learned K, Lee BT, Li CH, Raney BJ, Rhead B, Rosenbloom KR, Sloan CA, Speir ML, Zweig AS, Haussler D, Kuhn RM, Kent WJ. The UCSC genome browser database: 2014 update. Nucleic Acids Res. 2014; 42(Database issue):D764–D770.10.1093/nar/gkt1168 [PubMed: 24270787]
24. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci USA. 2005; 102(43):15545–15550.10.1073/pnas.0506580102 [PubMed: 16199517]
25. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G. Gene ontology: Tool for the unification of biology. Nat Genet. 2000; 25(1):25–29.10.1038/75556 [PubMed: 10802651]
26. Kanehisa M, Araki M, Goto S, Hattori M, Hirakawa M, Itoh M, Katayama T, Kawa-shima S, Okuda S, Tokimatsu T, Yamanishi Y. KEGG for linking genomes to life and the environment. Nucleic Acids Res. 2008; 36(S1):D480–D484.10.1093/nar/gkm882 [PubMed: 18077471]
27. Hosack DA, Dennis G Jr, Sherman BT, Lane HC, Lempicki RA. Identifying biological themes within lists of genes with EASE. Genome Biol. 2003; 4(10):R70.10.1186/gb-2003-4-10-r70 [PubMed: 14519205]
28. Serão NV, González-Peña D, Beever JE, Faulkner DB, Southey BR, Rodriguez-Zas SL. Single nucleotide polymorphisms and haplotypes associated with feed efficiency in beef cattle. BMC Genet. 2013; 14:94.10.1186/1471-2156-14-94 [PubMed: 24066663]
29. Delfino K, Rodriguez-Zas SL. Transcription factor-microRNA-target gene networks associated with ovarian cancer survival and recurrence. PLoS One. 2013; 8(3):e58608.10.1371/journal.pone.0058608 [PubMed: 23554906]
30. Hickman SE, Kingery ND, Ohsumi TK, Borowsky ML, Wan L, Means TK, Khoury JE. The microglial sensome revealed by direct RNA sequencing. Nat Neurosci. 2013; 16(12):1896–1905.10.1038/nn.3554 [PubMed: 24162652]

31. Martinez AN, Mehra S, Kaushal D. Role of Interleukin 6 in innate immunity to *Mycobacterium tuberculosis* infection. J Infect Dis. 2013; 207(8):1253–1261.10.1093/infdis/jit037 [PubMed: 23359591]
32. Sohn SH, Ko E, Kim SH, Kim Y, Shin M, Hong M, Bae H. Genome wide expression profile of *Agrimonia pilosa* in LPS-stimulated BV-2 microglial cells. Mol Cell Toxicol. 2009; 5(1):1–6.
33. Stables MJ, Shah S, Camon EB, Lovering RC, Newson J, Bystrom J, Farrow S, Gilroy DW. Transcriptomic analyses of murine resolution-phase macrophages. Blood. 2011; 118(26):e192–e208.10.1182/blood-2011-04-345330 [PubMed: 22012065]
34. Ramadas RA, Ewart SL, Medoff BD, LeVine AM. Interleukin-1 family member 9 stimulates chemokine production and neutrophil influx in mouse lungs. Am J Respir Cell Mol Biol. 2011; 44(2):134–145.10.1165/rcmb.2009-0315OC [PubMed: 20299540]
35. Magee DA, Taraktsoglou M, Killick KE, Nalpas NC, Browne JA, Park SD, Conlon KM, Lynn DJ, Hokamp K, Gordon SV, Gormley E, MacHugh DE. Global gene expression and systems biology analysis of bovine monocyte-dervied macrophages in response to in vitro challenge with *Mycobacterium bovis*. PLoS One. 2012; 7(2):e32034.10.1371/journal.pone.0032034 [PubMed: 22384131]
36. Ehrt S, Schnappinger D, Bekiranov S, Drenkow J, Shi S, Gingeras TR, Gaasterland T, Schoolnik G, Nathan C. Reprogramming of the macrophage transcriptome in response to interferon-γ and *Mycobacterium tuberculosis*: Signaling roles of nitric oxide synthase-2 and phagocyte oxidase. J Exp Med. 2001; 194(8):1123–1140.10.1084/jem.194.8.1123 [PubMed: 11602641]
37. MacMicking J, Xie QW, Nathan C. Nitric oxide and macrophage function. Annu Rev Immunol. 1997; 15(1):323–350.10.1146/annurev.immunol.15.1.323 [PubMed: 9143691]
38. Liu H, Liu Z, Chen J, Chen L, He X, Zheng R, Yang H, Song P, Weng D, Hu H, Fan L, Xiao H, Kaufmann SHE, Ernst J, Ge B. Induction of CCL8/MCP-2 by Mycobacteria through the activation of TLR2/PI3K/Akt signaling pathway. PLoS One. 2013; 8(2):e56815.10.1371/journal.pone.0056815 [PubMed: 23418602]
39. Kim BO, Liu Y, Zhou BY, He JJ. Induction of C chemokine XCL1 (lymphotactin/single C motif-1*α*/activation-induced, T cell-derived and chemokine-related cytokine) expression by HIV-1 tat protein. J Immunol. 2004; 172(3):1888–1895.10.4049/jimmunol.172.3.1888 [PubMed: 14734774]
40. de Oliveira CC, de Oliveira SM, Goes VM, Probst CM, Krieger MA, de Buchi DF. Gene expression profiling of macrophages following mice treatment with an immunomodulator medication. J Cell Biochem. 2008; 104(4):1364–1377.10.1002/jcb.21713 [PubMed: 18286468]
41. Mulcahy H, O'Rourke KP, Adams C, Molloy MG, O'Gara F. LST1 and NCR3 expression in autoimmune inflammation and in response to IFN-*γ*, LPS and microbial infection. Immunogenetics. 2006; 57(12):893–903.10.1007/s00251-005-0057-2 [PubMed: 16362817]
42. Osorio y Fortèa J, de La Llave E, Regnault B, Coppèe JY, Milon G, Lang T, Prina E. Transcriptional signatures of BALB/c mouse macrophages housing multiplying Leishmania amazonensis amastigotes. BMC Genomics. 2009; 10:119. [PubMed: 19302708]
43. Gundra UM, Girgis NM, Ruckerl D, Jenkins S, Ward LN, Kurtz ZD, Wiens KE, Tang MS, Basu-Roy U, Mansukhani A, Allen JE, Loke P. Alternatively activated macrophages derived from monocytes and tissue macrophages are phenotypically and functionally distinct. Blood. 201410.1182/blood-2013-08-520619
44. Uhrin P, Perkmann T, Binder B, Schabbauer G. ISG12 is a critical modulator of innate immune responses in murine models of sepsis. Immunobiology. 2013; 218(9):1207–1206.10.1016/j.imbio.2013.04.009 [PubMed: 23747037]
45. Pesce JT, Ramalingam TR, Wilson MS, Mentink-Kane MM, Thompson RW, Cheever AW, Urban JF Jr, Thomas A. Wynn TA. Retnla (Relma/Fizz1) suppresses Helminth-induced Th2-type-immunity. PLoS Pathog. 2009; 5(4):e1000393.10.1371/journal.ppat.1000393 [PubMed: 19381262]
46. Schmieder A, Michel J, Schönhaar K, Goerdt S, Schledzewski K. Differentiation and gene expression profile of tumor-associated macrophages. Semin Cancer Biol. 2012; 22(4):289–297.10.1016/j.semcancer.2012.02.002 [PubMed: 22349514]

47. Vogt L, Schmitz N, Kurrer MO, Bauer M, Hinton HI, Behnke S, Gatto D, Sebbel P, Beerli RR, Sonderegger I, Kopf M, Saudan P, Bachmann MF. VSIG4, a B7 family-related protein, is a negative regulator of T cell activation. J Clin Invest. 2006; 116(10):2817–2826.10.1172/JCI25673 [PubMed: 17016562]

48. Gladue DP, Zhu J, Holinka LG, Fernandez-Sainz I, Carrillo C, Prarat MV, O'Donnell V, Borca MV. Patterns of gene expression in swine macrophages infected with classical swine fever virus detected by microarray. Virus Res. 2010; 151(1):10–18.10.1016/j.vir-usres.2010.03.007 [PubMed: 20302897]

49. Moore KW, de Waal Malefyt R, Coffman RL, O'Garra A. Interleukin-10 and the interleukin-10 receptor. Annu Rev Immunol. 2001; 19:683–765.10.1146/annurev.immunol.19.1683 [PubMed: 11244051]

50. Sigruener A, Buechler C, Bared SM, Grandl M, Aslanidis C, Ugocsai P, Gehrmann M, Schmitz G. E-LDL upregulates TOSO expression and enhances the survival of human macrophages. Biochem Biophys Res Commun. 2007; 359(3):723–728.10.1016/j.bbrc.2007.05.169 [PubMed: 17553462]

51. Liu Q, Zheng J, Yin DD, Xiang J, He F, Wang YC, Liang L, Qin HY, Liu L, Liang YM, Han H. Monocyte to macrophage differentiation-associated (MMD) positively regulates ERK and Akt activation and TNF-a and NO production in macrophages. Mol Biol Rep. 2012; 39(5):5643–5650.10.1007/s11033-011-1370-5 [PubMed: 22203480]

52. Espejo C, Penkowa M, Demestre M, Montalban X, Martínez-Cáceres EM. Time-course expression of CNS inflammatory, neurodegenerative tissue repair markers and metallothioneins during experimental autoimmune encephalomyelitis. Neuroscience. 2005; 132(4):1135–1149.10.1016/j.neuroscience.2005.01.057 [PubMed: 16078373]

53. Karlstetter M, Walczak Y, Weigelt K, Ebert S, Van den Brulle J, Schwer H, Fuchshofer R, Langmann T. The novel activated microglia/macrophage WAP domain protein, AMWAP, acts as a counter-regulator of proinflammatory response. J Immunol. 2010; 185(6):3379–3390.10.4049/jimmunol.0903300 [PubMed: 20709948]

54. Trapnell C, Pachter L, Salzberg SL. TopHat: Discovering splice junctions with RNA-Seq. Bioinformatics. 2009; 25(9):1105–1111.10.1093/bioinformatics/btp120 [PubMed: 19289445]

55. Mende DR, Waller AS, Sunagawa S, Järvelin AI, Chan MM, Arumugam M, Raes J, Bork P. Assessment of metagenomic assembly using simulated next generation sequencing data. PLoS One. 2012; 7(2):e31386.10.1371/journal.pone.0031386 [PubMed: 22384016]

56. Ehrchen JM, Sunderkötter C, Foell D, Vogl T, Roth J. The endogenous Toll-like receptor 4 agonist S100A8/S100A9 (calprotectin) as innate amplifier of infection, autoimmunity, and cancer. J Leukoc Biol. 2009; 86(3):557–566.10.1189/jlb.1008647 [PubMed: 19451397]

57. Speranza V, Colone A, Cicconi R, Palmieri G, Giovanni D, Grassi M, Mattei M, Sali M, Delogu G, Andreola F, Colizzi V, Mariani F. Recombinant BCG-Rv1767 amount determines, in vivo, antigen-specific T cells location, frequency, and protective outcome. Microb Pathog. 2010; 48(5):150–159.10.1016/j.micpath.2010.02.003 [PubMed: 20219669]

58. Gordon S, Martinez FO. Alternative activation of macrophages: Mechanism and functions. Immunity. 2010; 32(5):593–604.10.1016/j.immuni.2010.05.007 [PubMed: 20510870]

59. Wu C, Pot C, Apetoh L, Thalhamer T, Zhu B, Murugaiyan G, Xiao S, Lee Y, Rangachari M, Yosef N, Kuchroo VK. Metallothioneins negatively regulate IL-27-induced type 1 regulatory T-cell differentiation. Proc Natl Acad Sci USA. 2013; 110(19):7802–7807.10.1073/pnas.1211776110 [PubMed: 23630250]

60. Inoue K, Takano H, Shimada A, Satoh M. Metallothionein as an anti-inflammatory mediator. Mediators Inflamml. 2009; 2009:101659.10.1155/2009/101659

61. Thamilarasan M, Hecker M, Goertsches RH, Paap BK, Schröder I, Koczan D, Thiesen HJ, Zettl UK. Glatiramer acetate treatment effects on gene expression in monocytes of multiple sclerosis patients. J Neuroinflammation. 2013; 10:126.10.1186/1742-2094-10-126 [PubMed: 24134771]

62. Dupèrè-Minier G, Desharnais P, Bernier J. Involvement of tyrosine phosphatase CD45 in apoptosis. Apoptosis. 2010; 15(1):1–13.10.1007/s10495-009-0413-z [PubMed: 19856105]

63. Lee HT, Kim SK, Kim SH, Kim K, Lim CH, Park J, Roh TY, Kim N, Chai YG. Transcription-related element gene expression pattern differs between microglia and macrophages during

inflammation. Inflamm Res. 2014; 63(5):389–397.10.1007/s00011-014-0711-y [PubMed: 24468891]

64. Przanowski P, Dabrowski M, Ellert-Miklaszewska A, Kloss M, Mieczkowski J, Kaza B, Ronowicz A, Hu F, Piotrowski A, Kettenmann H, Komorowski J, Kaminska B. The signal transducers Stat1 and Stat3 and their novel target Jmjd3 drive the expression of inflammatory genes in microglia. J Mol Med (Berl). 2014; 92(3):239–254.10.1007/s00109-013-1090-5 [PubMed: 24097101]
65. Mantovani A, Sica A, Sozzani S, Allavena P, Vecchi A, Locati M. The chemokine system in diverse forms of macrophage activation and polarization. Trends Immunol. 2004; 25(12):677–686.10.1016/j.it.2004.09.015 [PubMed: 15530839]
66. Marquis JF, Kapoustina O, Langlais D, Ruddy R, Dufour CR, Kim BH, MacMicking JD, Giguère V, Gros P. Interferon regulatory factor 8 regulates pathways for antigen presentation in myeloid cells and during tuberculosis. PLoS Genet. 2011; 7(6):e1002097.10.13/journal.pgen.1002097 [PubMed: 21731497]
67. Kota RS, Rutledge JC, Gohil K, Kumar A, Enelow RI, Ramana CV. Regulation of gene expression in RAW 264.7 macrophage cell line by interferon-gamma. Biochem Biophys Res Commun. 2006; 342(4):1137–1146.10.1016/j.bbrc.2006.02.087 [PubMed: 16516165]
68. Beutner C, Linnartz-Gerlach B, Schmidt SV, Beyer M, Mallmann MR, Staratschek-Jox A, Schultze JL, Neumann H. Unique transcriptome signature of mouse microglia. Glia. 2013; 61(9): 1429–1442.10.1002/glia.22524 [PubMed: 23832717]

Biographies

Scott E. Nixon is currently a Ph.D. Candidate in Bioinformatics in the laboratory of Dr. Rodriguez-Zas at the University of Illinois at Urbana-Champaign.

Dianelys González-Peña is currently a Ph.D. Candidate in Animal Genetics in the laboratory of Dr. Rodriguez-Zas at the University of Illinois at Urbana-Champaign.

Marcus A. Lawson is a Postdoctoral Researcher in the Division of Nutritional Sciences at the University of Illinois at Urbana-Champaign.

Robert H. McCusker is a Research Assistant Professor in the Department of Animal Sciences at the University of Illinois at Urbana-Champaign. His key areas of research include Neuroimmunology and the associated impacts on behavior.

Alvaro G. Hernandez is the Director of the High-Throughput Sequencing and Genotyping Unit, which is a part of the Roy J. Carver Biotechnology Center at the University of Illinois at Urbana-Champaign.

Jason C. O'Connor is an Assistant Professor in the Department of Pharmacology at the University of Texas Health Science Center in San Antonio. His key areas of research include Neuroinflammation and Inflammation-Associated Depression.

Robert Dantzer is a Professor in the Department of Symptom Research at the University of Texas MD Anderson Cancer Center in Houston. His key areas of research include Psychoneuroimmunology as well as Brain Expression and Action of Cytokines.

Keith W. Kelley is a Professor Emeritus in the Department of Animal Sciences at the University of Illinois at Urbana-Champaign. His key areas of research involve Immunophysiology and Behavior.

Sandra L. Rodriguez-Zas is a Professor in the Department of Animal Sciences, an Adjunct Professor in the Department of Statistics, an active member of the Neuroscience Program, and an affiliate with the Institute of Genomic Biology, all at the University of Illinois at Urbana-Champaign. Her key areas of research involve Statistical Genomics and Bioinformatics.

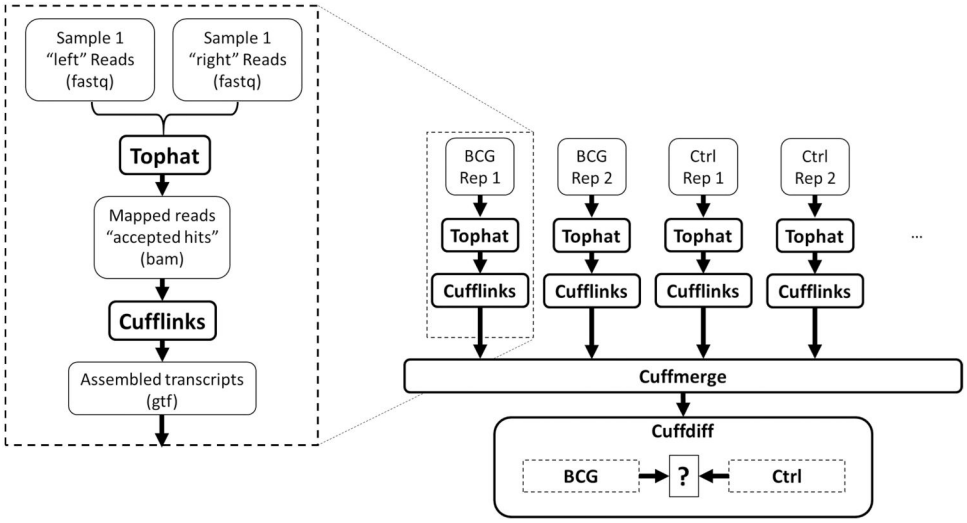

Fig. 1.

RNA-Seq workflow, showing the analysis of each sample individually by Tophat and Cufflinks (inset) before the collective analysis of all samples in Cuffdiff to test for differential expression (?) between conditions (BCG and Control or Ctrl groups).

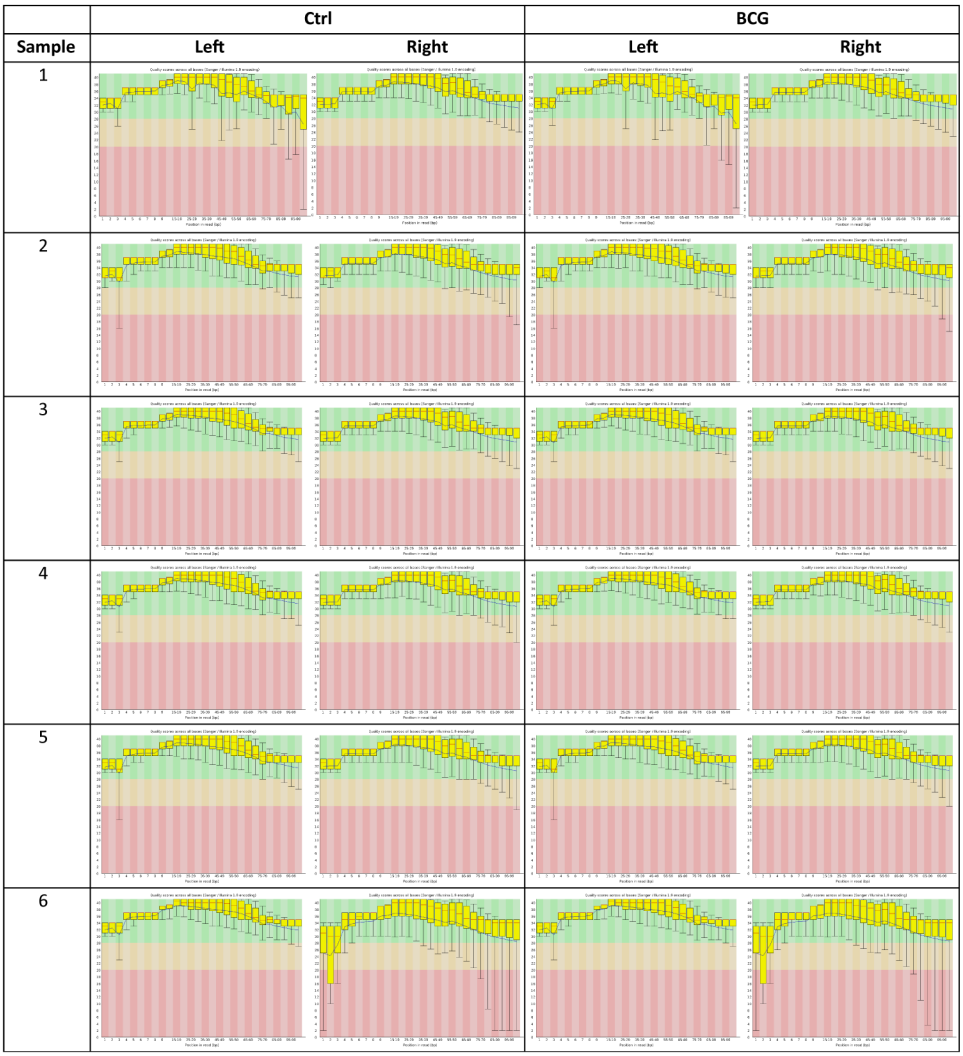

Fig. 2.

Quality box-and-whisker graphs via FastQC illustrating quality scores across the read length in the left and right reads from BCG and Control (Ctrl) samples.

Table 1

Read count of the paired-end (left/right) sample data.

Group	Sample	Reads[a]	Direction	Mapped (%)	Aligned pairs (%)
Ctrl	1	48,137,202	Left	78	77
		48,137,202	Right	81	
	2	54,465,049	Left	94	92
		54,465,049	Right	93	
	3	50,384,929	Left	95	94
		50,384,929	Right	95	
	4	58,747,069	Left	95	94
		58,747,069	Right	95	
	5	68,739,003	Left	95	93
		68,739,003	Right	94	
	6	45,314,712	Left	92	88
		45,314,712	Right	89	
BCG	1	64,092,232	Left	75	74
		64,092,232	Right	77	
	2	72,364,837	Left	94	92
		72,364,837	Right	94	
	3	67,524,327	Left	95	94
		67,524,327	Right	95	
	4	61,335,634	Left	95	94
		61,335,634	Right	95	
	5	54,744,239	Left	94	93
		54,744,239	Right	94	
	6	62,302,008	Left	92	89
		62,302,008	Right	90	

[a]Number of reads for the sample, for each of the paired runs (direction).

Table 2

Transcript and gene counts within Cuffdiff.

	Total tested	MA > 10[a]	Significant (FDR *p*-value < 0.05)	Named genes[b]
Transcript	62,490	29,844	2258	2079
Gene	23,274	12,009	1885	1884

[a] MA: Minimum alignment; a locus (i.e. transcript) needs at least this many fragments aligned before significance testing will be performed.

[b] Named genes were determined using the UCSC database (http://genome.ucsc.edu).

Table 3

Twenty five genes showing the greatest differential overexpression in BCG relative to Control (FDR p-value < 0.01).

Gene name[a]	Gene ID	Log_2 (BCG/Control)	Reference[b]
S100a9	20202	10.08	30
Ly6i	57248	9.50	31
Asprv1	67855	9.44	32, 33
Il1f9	215257	8.19	33, 34
Spon1	233744	8.02	35
Nos2	18126	8.01	36, 37
S100a8	20201	7.69	30, 36
Ccl8	20307	7.58	30, 38
Cxcr2	12765	7.50	30
Fcgr1	14129	6.03	30
Xcl1	16963	5.49	33, 39
AW112010	107350	4.98	40
Lst1	16988	4.80	41
Oas3	246727	4.77	67
Rsad2	58185	4.72	35
Smpdl3b	100340	4.52	42
Gbp2	14469	4.38	68
AA467197	433470	4.26	43
Ccl5	20304	3.99	33, 38
Cxcl10	15945	3.81	30
Chi3l3	12655	3.81	30
Acsl1	14081	3.59	35
Isg15	100038882	3.38	35
Ifi27l2a	76933	2.65	44
Arg1	11846	3.35	30, 33

[a] AA467197: expressed sequence AA467197; Acsl1: Acyl-CoA synthetase long-chain family member 1; Arg1: Arginase 1; Asprv1: Aspartic peptidase, retroviral-like 1; AW112010: Expressed sequence AW112010; Ccl5: Chemo-kine (C-C motif) ligand 5; Ccl8: Chemokine (C-C motif) ligand 8; Chi3l3: Chitinase-like 3; Cxcr2: Chemokine (C-X-C motif) receptor 2; Cxcl10: Che-mokine (C-X-C motif) ligand 10; Fcgr1: Fc receptor, IgG, high affinity I; Gbp2: Guanylate binding protein 2; Ifi27l2a: Interferon, alpha-inducible protein 27 like 2A; Il1f9: Interleukin 1 family, member 9; Isg15: ISG15 ubi-quitin-like modifier; Lst1: Leukocyte specific transcript 1; Ly6i: Lymphocyte antigen 6 complex, locus I; Nos2: Nitric oxide synthase 2, inducible; Oas3: 2′-5′ oligoadenylate synthetase 3; Rsad2: Radical S-adenosyl methionine domain containing 2; Smpdl3b: Sphingomyelin phosphodiesterase, acid-like 3B; Spon1: Spondin 1; S100a8: S100 calcium binding protein 8; S100a9: S100 calcium binding protein 9; Xcl1: Chemokine (C motif) ligand 1.

[b] Literature associating the listed gene with a macrophage population.

Table 4

Twenty five genes showing the greatest differential underexpression in the BCG relative to the control group (FDR p-value < 0.01).

Gene name[a]	Gene ID	Log_2 (BCG/Control)	Reference[b]
Retnla	57262	−4.90	30, 45
Cxcl13	55985	−4.67	30, 33, 35
Cd209a	170786	−4.42	33
Pf4	56744	−4.07	30, 35
Fcrls	80891	−3.90	30
Adm	11535	−3.51	30, 35
Lyve1	114332	−3.41	46
Vsig4	278180	−3.19	35, 47
Bank1	242248	−3.07	48
Il10	16153	−3.01	30, 33, 35, 49
Cd83	12522	−2.87	35
Faim3	69169	−2.79	50
Blk	12143	−2.79	35
Pou2af1	18985	−2.78	35
Mmd	67468	−2.71	35, 51
Cd79b	15985	−2.66	30
Bcar3	29815	−2.62	42
Cd2	12481	−2.55	35
Fabp4	11770	−2.43	30, 35
Gimap6	231931	−2.36	35
F13a1	74145	−2.35	30
Ptprcap	19265	−2.34	—
Phgdh	236539	−1.94	43
Mt1	17748	−1.70	52
Wfdc17	100034251	−1.22	53

[a] Adm: Adrenomedullin; Bank1: B cell scaffold protein with ankyrin repeats 1; Bcar3: Breast cancer anti-estrogen resistance 3; Blk: B lymphoid kinase; Cd2: Cd2 antigen; Cd209a: Cd209a antigen; Cd79b: Cd79b antigen; Cd83: CD83 antigen; Cxcl13: Chemokine (C-X-C motif) ligand 13; Fabp4: Fatty acid binding protein 4, adipocyte; Faim3: Fas apoptotic inhibitory molecule 3; Fcrls: Fc receptor-like S, scavenger receptor; F13a1: Coagulation factor XIII, A1 subunit; Gimap6: GTPase, IMAP family member 6; Il10: Interleukin 10; Lyve1: Lymphatic vessel endothelial hyaluronan receptor 1; Mmd: Monocyte to macrophage differentiation-associated; Mt1: Metallothionein 1; Pf4: Platelet factor 4; Phgdh: 3-phosphoglycerate dehydrogenase; Pou2af1: POU domain, class 2, associating factor 1; Ptprcap: Protein tyrosine phosphatase, receptor type, C polypeptide-associated protein; Retnla: Resistin like alpha; Wfdc17: WAP four-disulfide core domain 17; Vsig4: V-set and immunoglobulin domain containing 4.

[b] Literature associating the listed gene with a macrophage population.

Table 5

List of member terms for each functional clusters (ES > 4).

Cluster identifier	ES (Genes)[a]	Identifier (Genes)[b]	Term name
1	17.89 (98)	GO:0006412 (69)	Translation
		GO:0003735 (59)	Structural constituent of ribosome
		GO:0005198 (85)	Structural molecule activity
		Mmu03010 (60)	Ribosome
2	11.24 (118)	GO:0009611 (80)	Response to wounding
		GO:0006952 (92)	Defense response
		GO:0006954 (54)	Inflammatory response
3	7.15 (66)	GO:0042330 (35)	Taxis
		GO:0006935 (35)	Chemotaxis
		GO:0007626 (46)	Locomotory behavior
		GO:0008009 (17)	Chemokine activity
		GO:0042379 (17)	Chemokine receptor binding
		GO:0005125 (38)	Cytokine activity
4	5.89 (67)	GO:0030246 (67)	Carbohydrate binding
		GO:0030247 (30)	Polysaccharide binding
		GO:0001871 (30)	Pattern binding
		GO:0005539 (27)	Glycosaminoglycan binding
5	5.83 (83)	GO:0006915 (79)	Apoptosis
		GO:0012501 (79)	Programmed cell death
		GO:0008219 (83)	Cell death
		GO:0016265 (83)	Death
6	5.45 (88)	GO:0042981 (88)	Regulation of apoptosis
		GO:0043067 (88)	Regulation of programmed cell death
		GO:0010941 (88)	Regulation of cell death

[a] ES: Enrichment Score; Listed in parenthesis, the ES is the number of genes enriching the cluster.

[b] Member terms of the cluster, with the number of genes enriching that term in parentheses.