A new technique for obtaining whole pathogen transcriptomes from infected host tissues

Tatyana L. Azhikina1, Timofey A. Skvortsov1, Tatyana V. Radaeva2, Andrey V. Mardanov3, Nikolay V. Ravin3, Alexander S. Apt2, and Eugene D. Sverdlov1

1Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, Moscow, Russia, 2Central Institute for Tuberculosis, Moscow, Russia, and 3Centre “Bioengineering,” Russian Academy of Sciences, Moscow, Russia

Material and Methods

Basic Protocols

Growth and transformation of Escherichia coli, preparation of plasmid DNA, gel electrophoresis, and other standard manipulations were performed as described (1). M. tuberculosis H37Rv genomic DNA was isolated according to Reference 2.

Mice and infection

Mice of the A/SnCit strain (hereafter abbreviated as A/Sn) were bred under conventional conditions at the Animal Facilities of the Central Institute for Tuberculosis (CIT; Moscow, Russia), in accordance with the guidelines from the Russian Ministry of Health #755, National Institutes of Health (NIH) Office of Laboratory Animal Welfare (OLAW) Assurance #A5502–01. Water and food were provided ad libitum. Female mice 2.5 months of age were infected via the respiratory route using an Inhalation Exposure System (Glas-Col, Terre Haute, IN, USA). The animals were exposed for 40 min to an aerosol produced by nebulizing 8 mL bacterial suspension in 50-μL volume containing 10 pM of each primer (Supplementary Table S2) and 1 U Taq DNA recombinant polymerase (MBI Fermentas, Vilnius, Lithuania). After 28 cycles of PCR, 5 μL reaction mixture was taken away every two cycles, and amplification continued up to the 40th cycle. The amplification products were analyzed by gel electrophoresis in 2.0% agarose gel. Expression levels were estimated by the difference in number of cycles after which the product could be visualized (~20 ng of the product in an aliquot). To exclude contamination by genomic DNA, RT-experiments were done in parallel. All RT-PCR experiments were reproduced at least three times with independent cDNA preparations.

454 pyrosequencing

Nucleotide sequences of cDNAs were determined by massively parallel pyrosequencing using a GS FLX genetic analyzer (Roche, Mannheim, Germany) and a 25 × 75 cm pictoriter plate. The sequences were read in 98,692 independent reactions, which correspond to a total of 20,193,290 nucleotides. The number of cDNA sequences complementary to each gene was found using BLASTn: a sequence was assigned to a particular gene if it was >95% complementary to a gene segment longer than 40 nucleotides.

Mycobacterium tuberculosis whole-transcriptome features

The transcribed genes were grouped using a functional classification tree (http://genolist.pasteur.fr/Tuberculist). The representation in each category was determined, which appeared to be different from that in the genome (Supplementary Table S5). Due to the close similarity of transposase genes, we were unable to unambiguously assign cDNA fragments to the genes of this functional category (insertion sequences and phages), and it appeared that the expression of the transposase family genes encoding methoxy mycolic acid synthases. Up-regulation of some of these genes in the mouse phagosome progression (e.g., fadD28 and fadD26) was previously reported (3,4).

Among the expressed genes (Supplementary Table S5), the genes involved in lipid metabolism were actively transcribed (49 of 226 genes identified in this pathway were expressed). In particular, we revealed the expression of genes of fatty acid biosynthesis pathway: 9 of 34 genes belonging to the fadD family and encoding acyl-CoA dehydrogenases, indicating both lipolytic and lipogenic activities of mycobacteria in the infected lung. Three of five ppsA–ppsE family genes were expressed. The products of these genes are responsible for the synthesis of phthiocerol, a progenitor of phthiocerol dimycocerosates (PDIMs), which are involved in the mycobacterial cell wall building. Genes involved in the synthesis and export of surface-exposed lipid virulence factors, such as PDIM and sulfolipid-1 (SL-1), are required for bacterial growth in mice (5). In addition, surface-exposed lipids provide protection against host-induced damage, as well as modulate the immune response to infection (6).

Three out of four mmaA1–mmaA4 family genes encoding mycolic acid synthases were expressed in our system. Induction of these genes in the mouse phagosome was reported earlier and suggested to be linked to the maintenance of cell wall hydrophobicity (7). The M. tuberculosis genome contains a cluster of polyketide synthase genes (pks1–pks17), whose products are thought to be involved in the synthesis of phthiocerol and phenylphthiocerol. In our method, 6 of 16 genes from this family were expressed. Induction of pks genes during M. tuberculosis persistence within macrophages was
repeatedly reported by different authors (8,9). The family of polyketide synthase-associated proteins, PapA1–Pap5, is extremely important for mycobacterial survival (10). Our results demonstrate orchestrated expression of these genes, which is in agreement with their genomic organization as functional operons that also contain pkts and mmpL (membrane protein) families.

The spectrum of transcribed genes from the respiratory metabolism category (Supplementary Table S3) indicates a switching from aerobic to anaerobic respiration, as described earlier for the chronic phase of infection (11). Thus, the transcription of the narH, narI, and narf genes from the narGHJI operon that encodes nitrate reductases—as well as narX and narK3 encoding nitrate ion transporters—was noticed. On the other hand, a characteristic decrease in the transcription of genes for various ATPases occurred: only atpD of the atpA–atpH gene cluster was found to be transcribed.

References


